

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 8, 2003, 15:31:12 ; Search time 7863 Seconds
(without alignments)
11636.681 Million cell updates/sec

Title: US-09-831-846-1

Perfect score: 3144

Sequence: 1 gcctggctccctctcgtgag.....gcgggtttggtgattgcttatg 3144

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_htg.*

3: gb_in.*

4: gb_om.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl.*

9: gb_pr.*

10: gb_ro.*

11: gb_sts.*

12: gb_sy.*

13: gb_un.*

14: gb_vi.*

15: em_ba.*

16: em_fun.*

17: em_hum.*

18: em_in.*

19: em_lm.*

20: em_om.*

21: em_or.*

22: em_ov.*

23: em_pat.*

24: em_ph.*

25: em_pl.*

26: em_ro.*

27: em_sts.*

28: em_un.*

29: em_vi.*

30: em_htg_hum.*

31: em_htg_inv.*

32: em_htg_other.*

33: em_htg_mus.*

34: em_htg_pln.*

35: em_htg_rod.*

36: em_htg_man.*

37: em_htg_vrt.*

38: em_sy.*

39: em_htgo_hum.*

40: em_htgo_mus.*

41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3144	100.0	3144	9	AB033072
2	2927	93.1	3320	9	AB056799
C	1424.8	45.3	100426	9	AL591063
3	11282	40.8	179820	2	AC011679
4	1166.6	37.1	159443	2	AC101792
C	1166.6	37.1	303563	2	AC124583
C	1166.6	37.1	308013	2	AC125047
8	1151.4	36.6	197893	2	AC127880
C	827.4	26.3	197893	2	AC127880
10	634.6	20.2	1887	6	AX179258
11	634.6	20.2	2549	9	BC003578
12	633	20.1	1887	6	AX266997
13	633	20.1	2558	6	AX266995
14	621	19.8	2458	9	BC015581
15	608	19.3	1653	6	AX358336
16	587.8	18.7	2612	10	BC023156
17	563.8	17.9	3174	9	AB040917
18	554.6	17.6	1653	6	AX358338
19	549.2	17.5	3047	9	AK055365
20	544.4	17.3	3701	9	AK096627
C	516.4	16.4	180153	2	AC018798
21	516.4	16.4	180153	2	AC018798
22	511	16.3	139072	9	AP000485
23	511	16.3	182769	9	AP003176
24	477.2	15.2	92797	9	AF038458
C	472	15.0	138538	9	AC011445
C	472	15.0	172090	2	AC018477
27	465	14.8	71605	2	AC119332
28	462.6	14.7	1889	9	BC027475
C	447.8	14.2	173649	2	AC117169
C	447.6	14.2	118112	9	HS462C17
31	446.4	14.2	1683	10	BC004018
32	438	13.9	168508	2	AC087141
33	438	13.9	198578	2	AC087138
C	432.6	13.8	279589	2	AC073768
C	403.6	12.8	158885	9	CNS01DWT
36	401.4	12.8	4301	9	HS803436
37	346.6	11.0	1832	9	BC025310
C	308.6	9.8	138034	2	AC096476
C	305.6	9.7	205026	2	AC106165
40	263.6	8.4	2271	9	BC014678
C	252.2	8.0	100660	2	AC095875
C	250.4	8.0	63949	2	AC110498
43	213.2	6.8	1307	9	BC014040
44	208	6.6	1286	9	BC000207
45	204.2	6.5	432	6	AX358333

ALIGNMENTS

RESULT 1
AB033072

LOCUS

DEFINITION

AB033072

ACCESSION

VERSION

KEYWORDS

SOURCE

AB033072 3144 bp mRNA linear PRI 11-NOV-1999
Homo sapiens mRNA for KIAA1246 protein, partial cds.

AB033072.1 GI:6330832

Homo sapiens brain cDNA to mRNA, clone_lib:pBluescriptII SK plus

Clone:hh00149a.

Homo sapiens

ORGANISM

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE

1 (sites)

AUTHORS

Nagase,T., Ishikawa,K., Kikuno,R., Hirotsawa,M., Nomura,N. and Ohara,O.

1441 GATGACCGGCTGGTAGGAACTCTCAAGGACCGTGTCTATGACAAATGCAACCCCTGGAC 1500
 1441 GATGACCGGCTGGTAGGAACTCTCAAGGACCGTGTCTATGACAAATGCAACCCCTGGAC 1500
 1501 ATCTTCATACCAACATCTCAGACAGTGTGTCCTTACCTGCAATGCTGCAATGCTGCC 1560
 1501 ATCTTCATACCAACATCTCAGACAGTGTGTCCTTACCTGCAATGCTGCAATGCTGCC 1560
 1561 GGAGAGCCACGGCCATGTTGAGGTTCTCATGCTCCAGTCCCAACCTCAGCAACAGC 1620
 1561 GGAGAGCCACGGCCATGTTGAGGTTCTCATGCTCCAGTCCCAACCTCAGCAACAGC 1620
 1621 ACCAGCCGACTGCACCCCTCAAGTCCCGCTCTCAGACATCACTGGCTCCAGCAAGCC 1680
 1621 ACCAGCCGACTGCACCCCTCAAGTCCCGCTCTCAGACATCACTGGCTCCAGCAAGCC 1680
 1681 AGCCGGGAGTGGAGGAGTGGGGGCGAGAGCTCCCAAAAGCCCGGNAACGGGCT 1740
 1681 AGCCGGGAGTGGAGGAGTGGGGGCGAGAGCTCCCAAAAGCCCGGNAACGGGCT 1740
 1741 GTGCTGTGTCGAGTGAACACCACTCGGCCCTGTCGAGTGGTCAAGTGGTCTGTCAGCAAGTCA 1800
 1741 GTGCTGTGTCGAGTGAACACCACTCGGCCCTGTCGAGTGGTCAAGTGGTCTGTCAGCAAGTCA 1800
 1801 GCACCCCGGTTGAAGTGTACCAAGTGCAGTGCAGTCAACTGCTCTGACGATGAGTACTGATT 1860
 1801 GCACCCCGGTTGAAGTGTACCAAGTGCAGTGCAGTCAACTGCTCTGACGATGAGTACTGATT 1860
 1861 TACAGGATATCCAGCTCCAAAGAGGCTTGGTGTCAACAACTGGTGTGTCAGGGACT 1920
 1861 TACAGGATATCCAGCTCCAAAGAGGCTTGGTGTCAACAACTGGTGTGTCAGGGACT 1920
 1921 GGTACGACTTGTGTTGGCCATGTTGGATGACACAGCCACGACACTCAGCGGCCACC 1980
 1921 GGTACGACTTGTGTTGGCCATGTTGGATGACACAGCCACGACACTCAGCGGCCACC 1980
 1981 AACATGTTGGCTGCGCCAGTTCACCAAGGCTGACTACCCGAGTGCACAGTCCATG 2040
 1981 AACATGTTGGCTGCGCCAGTTCACCAAGGCTGACTACCCGAGTGCACAGTCCATG 2040
 2041 CACAGCAGATCTGGGGGACCATGATCTGGTCAATCGGGGCAATCGTGGCCAGC 2100
 2041 CACAGCAGATCTGGGGGACCATGATCTGGTCAATCGGGGCAATCGTGGCCAGC 2100
 2101 CTGCTGTTCTCATCGTCACTCTCATGTCGCTCAAGGTTGGTCAAGGTTGGCAACCCAGGGCCCC 2160
 2101 CTGCTGTTCTCATCGTCACTCTCATGTCGCTCAAGGTTGGTCAAGGTTGGCAACCCAGGGCCCC 2160
 2161 AGCAAGATGGCAGCGCCGTGAGCAATGTGTAATCGCAGACCAACCGCCGCGCCAGCCCG 2220
 2161 AGCAAGATGGCAGCGCCGTGAGCAATGTGTAATCGCAGACCAACCGCCGCGCCAGCCCG 2220
 2221 CCTCCAGCAGGACACAGCGGGGCGCCGCGCAGGGGCGCCGCAAGTGTGTTGGCC 2280
 2221 CCTCCAGCAGGACACAGCGGGGCGCCGCGCAGGGGCGCCGCAAGTGTGTTGGCC 2280
 2281 AACGAGCTCTGACTTACCGCAGCTGGCCGCGCCGCGCAGTCACTCTTCCCTCCAGC 2340
 2281 AACGAGCTCTGACTTACCGCAGCTGGCCGCGCCGCGCAGTCACTCTTCCCTCCAGC 2340
 2341 TCCCTGGGAGTGGGAGGCTGGGGGCTGGGACGGGCGCCCTGGAGGATCCACCCCTCC 2400
 2341 TCCCTGGGAGTGGGAGGCTGGGGGCTGGGACGGGCGCCCTGGAGGATCCACCCCTCC 2400
 2401 GCCCGGCGCCCAAGCCAGCTTGAACCGCTGATGGGGCTTCCGCTCCCTGGACCTC 2460
 2401 GCCCGGCGCCCAAGCCAGCTTGAACCGCTGATGGGGCTTCCGCTCCCTGGACCTC 2460
 2461 AAGAGTCAGAAAGAGGAGTGTGGACTCCAGACTCCAGACTCCAGCGCGGGAGGGGCTGG 2520
 2461 AAGAGTCAGAAAGAGGAGTGTGGACTCCAGACTCCAGACTCCAGCGCGGGAGGGGCTGG 2520

2521 ACGTCGGCCCGGGCCACACCTCGAGCCGAGAGCCTGCTGGGGCCCCCTGGGGCCCCG 2580
 2521 ACGTCGGCCCGGGCCACACCTCGAGCCGAGAGCCTGCTGGGGCCCCCTGGGGCCCCG 2580
 2581 GCCAGAGCTCTCCCTTGGAGGGCAAGGCAAAAGCCAGCCACTCCCTTCGAC 2640
 2581 GCCAGAGCTCTCCCTTGGAGGGCAAGGCAAAAGCCAGCCACTCCCTTCGAC 2640
 2641 ATGGGGGACTTTCGCTGCTCGGGCGGGAGGGGCTGTCGCCGGGGGCTTACAGTCCCT 2700
 2641 ATGGGGGACTTTCGCTGCTCGGGCGGGAGGGGCTGTCGCCGGGGGCTTACAGTCCCT 2700
 2701 CGGAAGTTCGAAATCTGGACGAAAGCGAGCCCTCTCTGCAACGGCATGCTCTTGGCC 2760
 2701 CGGAAGTTCGAAATCTGGACGAAAGCGAGCCCTCTCTGCAACGGCATGCTCTTGGCC 2760
 2761 TTTGAGGAGTGAAGTGGTGGGGGCGGGGAGCTTTTGGCAGCTCCGAAATGGGTGATG 2820
 2761 TTTGAGGAGTGAAGTGGTGGGGGCGGGGAGCTTTTGGCAGCTCCGAAATGGGTGATG 2820
 2821 GAGAGCAGCTTAGTGGGGTGGGCATGCTCCCTTCTGTCGCGAGGGTGGGAGAAG 2880
 2821 GAGAGCAGCTTAGTGGGGTGGGCATGCTCCCTTCTGTCGCGAGGGTGGGAGAAG 2880
 2881 GGAAGAAATCTCACTGGCAAGTGTGAGTGTTCATGTGATGTTTACATCCAGGG 2940
 2881 GGAAGAAATCTCACTGGCAAGTGTGAGTGTTCATGTGATGTTTACATCCAGGG 2940
 2941 ACAGTTCGTCCTCCCTGTCGCTGTCGCTGTCGCTGTCGCTGTCGCTGTCGCTGTCGCT 3000
 2941 ACAGTTCGTCCTCCCTGTCGCTGTCGCTGTCGCTGTCGCTGTCGCTGTCGCTGTCGCT 3000
 3001 CCTCCCGCAGCCCGGGGCTGCTCAGGAAATGTGGACTCGCTCAAATCCAGGACT 3060
 3001 CCTCCCGCAGCCCGGGGCTGCTCAGGAAATGTGGACTCGCTCAAATCCAGGACT 3060
 3061 GAGCCCTGAGTGTGGAAAGCGGAGCTCCGCCCTTCTAAATCACAATGTAGCCTACAA 3120
 3061 GAGCCCTGAGTGTGGAAAGCGGAGCTCCGCCCTTCTAAATCACAATGTAGCCTACAA 3120
 3121 GCAAGCGCTTGGATTGCTTATG 3144
 3121 GCAAGCGCTTGGATTGCTTATG 3144

RESULT 2
 AB056799
 LOCUS AB056799
 DEFINITION Macaca fascicularis brain cDNA clone:Qf1A-11865, full insert sequence.
 ACCESSION AB056799
 VERSION GI:13365898
 KEYWORDS fis (full insert sequence); oligo capping.
 SOURCE Macaca fascicularis adult male frontal lobe left cDNA to mRNA, clone_lib:macaque brain cDNA library Qf1A clone:Qf1A-11865.
 ORGANISM Macaca fascicularis
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea; Cercopitheciinae; Macaca.
 REFERENCE 1 (sites)
 AUTHORS Osada, N., Hida, M., Kusuda, J., Tanuma, R., Iseki, K., Hirai, M., Terao, K., Suzuki, Y., Sugano, S. and Hashimoto, K.
 TITLE Isolation of full-length cDNA clones from macaque brain cDNA libraries
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 3120)
 AUTHORS Hashimoto, K., Osada, N., Hida, M., Kusuda, J. and Sugano, S.
 TITLE Direct Submission
 JOURNAL Submitted (09-MAR-2001) Katsuyuki Hashimoto, National Institute of Infectious Diseases, Division of Genetic Resources, 23-1, Toyama 1-chome, Shinjuku-ku, Tokyo 162-8640, Japan
 (E-mail: khashi@nih.go.jp, URL: http://www.nih.go.jp/yoken/genebank/, Tel: 81-3-5285-1111 (ex. 2120), Fax: 81-3-5285-1181)

299	Qy	CATCTGCTGCATT-----CGCCGGAGCCCGCGGGTTAGCTCG	340
379	Db	CACTGCTGCGTTTCAGAGCCCGCGGAGCCCGCGGGAGGTTTAGCTCG	438
341	Qy	AGTCTGTCTGGGGGGAGGAGTGGTGGCGAGCGGGGAGCCCGGGCCCGCGGA	400
439	Db	AGTCTGTCTGGGGGGAGGAGTGGTGGCGAGCGGGGAGCCCGGGCCCGCGGA	498
401	Qy	CGCGGCTCGGTGCCACCCAGCCGGGGTAGATGCTGCCTCGCCAGCCGCTG	453
499	Db	CGCGGCTCGGTGCCACCCAGCCGGGGTAGATGCTGCCTCGCCAGCCGCTG	558
454	Qy	AGTGACAGACCAATGGAGACCCCTGCTTGGTGGCTGCTAGCGTTGGCATGG	513
559	Db	AGTGACAGACCAATGGAGACCCCTGCTTGGTGGCTGCTAGCGTTGGCATGG	618
514	Qy	GTGTCAGAGCCCTGCCCAAGTACTGTGTCGCAGAACTGCTGAGTCACTGGG	573
619	Db	GTGTCAGAGCCCTGCCCAAGTACTGTGTCGCAGAACTGCTGAGTCACTGGG	678
574	Qy	CTGTGCCCTCCAAAGGGGCTCTTTTGTACCCCTGATATTGACCGGGGACAG	633
679	Db	CTGTGCCCTCCAAAGGGGCTCTTTTGTACCCCTGATATTGACCGGGGACAG	738
634	Qy	CTGGCCCTGGGGGCAACTTCATCATCAGCCGCGAGGACTTTGCCAATGACG	693
739	Db	CTGGCCCTGGGGGCAACTTCATCATCAGCCGCGAGGACTTTGCCAATGACG	798
694	Qy	GGGCTGTGGACTGACCCCTGCCAGGAACACCATCAGCCATCCAGCCCTTTT	753
799	Db	GGGCTGTGGACTGACCCCTGCCAGGAACACCATCAGCCATCCAGCCCTTTT	858
754	Qy	CTGGACCTCGAGAGCTCCGCTCCCTGATCTTGAAGAAATCGGCTGCGCAAG	813
859	Db	CTGGACCTCGAGAGCTCCGCTCCCTGATCTTGAAGAAATCGGCTGCGCAAG	918
814	Qy	GAGGACACTCCCGGGGCTTCAACTGAGGACCTTTCGTGAACAAACAACAGCTG	873
919	Db	GAGGACACTCCCGGGGCTTCAACTGAGGACCTTTCGTGAACAAACAACAGCTG	978
874	Qy	GCGGCATCGGAGATGAGGCTTTTTCAGGACTTCTGTCGATGAGGATCTGGAC	933
979	Db	GCGGCATCGGAGATGAGGCTTTTTCAGGACTTCTGTCGATGAGGATCTGGAC	1038
934	Qy	TCCCTAAACAACCTCCATGGCTCCGCTGGGACTCCGTCGCGGCGGCTCAAC	993
1039	Db	TCCCTAAACAACCTCCATGGCTCCGCTGGGACTCCGTCGCGGCGGCTCAAC	1098
994	Qy	CAGCTGAGCTGAGCACAACTGCTGGATCACATCGCCGGGCGACCTTTTCAGAC	1053
1099	Db	CAGCTGAGCTGAGCACAACTGCTGGATCACATCGCCGGGCGACCTTTTCAGAC	1158
1054	Qy	CAGAACTGGCCCGCTGGATCTCACTCAATCGGCTGCAAGTGGCCCCCTGATCC	1113
1159	Db	CAGAACTGGCCCGCTGGATCTCACTCAATCGGCTGCAAGTGGCCCCCTGATCC	1218
1114	Qy	ATCTTTTGGCTCCAGGCTTGGCTTTCAGGACCACTTTGCCCCACCTTTGTC	1173
1219	Db	ATCTTTTGGCTCCAGGCTTGGCTTTCAGGACCACTTTGCCCCACCTTTGTC	1278
1174	Qy	TTTAGTTTGGGGTAAACCACCTTCACTCAATTTGAGCTTCTCTGGTGGGAGG	1233
1279	Db	TTTAGTTTGGGGTAAACCACCTTCACTCAATTTGAGCTTCTCTGGTGGGAGG	1338
1234	Qy	GAGCGGAGCGACTGCTGAAACCTTGGCTCCAGGGGGCTCAAGGGTCTGCTACT	1293
1339	Db	GAGCGGAGCGACTGCTGAAACCTTGGCTCCAGGGGGCTCAAGGGTCTGCTACT	1398
1294	Qy	TGSCATGTGGCTGAGGAGGTTGTGCGAGCCGCTCTCATACCCAGCACACAC	1353
1399	Db	TGSCATGTGGCTGAGGAGGTTGTGCGAGCCGCTCTCATACCCAGCACACAC	1458
1354	Qy	AAGTTGCTGTTCTTGGAGGGCCAGGCGGCGACACTCAAGTGTCAAGCAATTTGGG	1413

Lab host: TOP10
 Vector: pME18S-FL3 (Acc. No. AB009864)
 R. Site1: DraIII (CACTGTGTG)
 R. Site2: DraIII (CACCATGTG)
 Description: 1st strand cDNA was primed with an oligo (dT) primer [ATGTGGCTTTTGTTTT] double-stranded cDNA was synthesized using specific 5' and 3' primers and amplified by PCR. The PCR product was digested with SfiI and size selection was performed to exclude fragments <1.5kb. The SfiI-digested PCR product was cloned into distinct DraIII sites of pME18S-FL3. XhoI sites just outside the DraIII sites can be used to isolate the cDNA insert. Libraries were constructed by Sugano et al. (University of Tokyo, Institute of Medical Science). Custom primer used for sequencing (5' end primer [CTTCTGCTTAAAGCTGG]; 3' end primer [CGACCTGCAGCTCGCAC]).

Location/Qualifiers
 1..3320
 /organism="Macaca fascicularis"
 /db_xref="taxon:9541"
 /clone="Qf1A-11865"
 /sex="male"
 /tissue type="frontal lobe left"
 /clone_lib="macaque brain cDNA library Qf1A"
 /dev_stage="adult"
 /notes="Host: TOP10; Vector: pME18S-FL3 (Acc. No. AB009864); R. Site1: DraIII (CACTGTGTG); R. Site2: DraIII (CACCATGTG)"
 571..2940

CDS
 /codon_start=1
 /product="hypothetical protein"
 /protein_id="BAB39323.1"
 /db_xref="GI:13365899"
 /translation="METLGGLLAFGMFAVVDACPKYCVCONLSBSLGLCPKSKLLRSPFDLDRTRVELRGLGNFLIHSRQDFANMGLVLDLTLRSNTLSHIQPFSLDELSLHLSLDSNRLPGEDETLRVLNHLI VNNQGLGDEAFEDFLELDDLSLNNLHGLPVSVMWNLHSLDHLNHLHIAETGADLQKRLDLTNSRLOKLPDPIF ARSOALATPAPLPSPFSGNPLHCNCELMLRRLRDLDTGSPFGLKGRYF WHYREDFCEPPLI TOHTHKLVLVGOAATLCKAIGDPSPLIHWVAPDRLVNSG RTAVDNGTLDLITTSQDSGACTCAANAAAGATAVEVSI VQLPHLSNSTRAPP KRLSLDITGSSKTSRGGGGGEPKSPERAVLIVSEVTISALAKMSVSKSTPRV MYQLQNSDDEVLIYRMI PANKAFVNNLVSGTVDLCLVLANWDDTATLTLNINP GCAOFFTKADYPOCSNHSQILGGTMIIVGIVATLVLLVILMRYKVCNHNAPS KMAAVNSVYVOTNGAPPPSPAPAGAPQPPKVVNRNELLDFASLARASDSSS SLLGSGEAGLGRAPWLPSPAPRPSLDRLMGAFASLDLKSQRKEILLDRTPAGR GAGTSARHRSRDLPLGPPAARARSLPLLEGKAKRSHSFDMDGFAAAAGGVVPG GYSPPRVSNIVWIKRSUVNGLMLLPFEESDLVNGARTFGSSEWMBSTV"

BASE COUNT 628 a 1071 c 967 g 654 t

Query Match 93.1%; Score 2927; DB 9; Length 3320;
 Best Local Similarity 96.5%; Pred. No. 0;
 Matches 3068; Conservative 0; Mismatches 75; Indels 36; Gaps 6;

Qy	1	GCCTGGCTCCCTCGCTGAGACACATACACTCACATACACACCCCGGAGGCTCG	60
Db	80	GCCTGGCTCCCTCGCTGAGACACATACACTCACATACACACCCCGGAGGCTCG	138
Qy	61	TCTGAACTGAAGACACCCCATTCCAAGATGCCGAGTTCCTGGGAATGCCTGGGT	120
Db	139	TCTGAACTGAAGACACCCCATTCCAAGATGCCGAGTTCCTGGGAATGCCTGGGT	198
Qy	121	TCTTCGATCCGGAAATCCTACCGGCATCCCTCCTAGGAGGATATTATTATTAT	178
Db	199	TCTTCGATCCGGAAATCCTACCGGCATCCCTCCTAGGAGGATATTATTATTAT	258
Qy	179	TCTTTAATCTGGAAGAGAGAA CAAGTTGTGCTTTTCCCGCTTCTTCTGCTAAC	238
Db	259	TCTTTAATCTGGAAGAGAGAA CAAGTTGTGCTTTTCCCGCTTCTTCTGCTAAC	318
Qy	239	GCCNTGGATTAACGAATAAGCGCTCAGGCTTTCCCGCTTCTCAGGCTCCGAGCC	298
Db	319	GCCNTGGATTAACGAATAAGCGCTCAGGCTTTCCCGCTTCTCAGGCTCCGAGCC	378

1459	Db		AAAGTGTCTAGTTCTGGAGGGCCAGGCGGCACACTCAAGTGTCAAAGCCATCGGGGACCCC	1518
1414	QY		AGCCCCCTTATCCACTGGGTAGCCCCCGGATGACCCGCTGGTAGGGAATCTCTCAAGGACC	1473
1519	Db		AGCCCCCTTATCCACTGGGTAGCCCCCGGATGACCCGCTGGTAGGGAATCTCTCAAGGACT	1578
1474	QY		GCTGTCTATGACAAATGGACCCCTGGACATCTTTCATCACACATCTCAGACAGTGTGGCC	1533
1579	Db		GCTGTCTATGACAAATGGACCCCTGGACATCTTTCATCACACATCTCAGACAGTGTGGCC	1638
1534	QY		TTACACTGTATGCTGCTGCTGCTGCGGAGAGCCACCGCCATGCTGAGTGTCCATC	1593
1639	Db		TTACACTGTATGCTGCTGCTGCTGCGGAGAGCCACCGCCATGCTGAGTGTCCATC	1698
1594	QY		GTCCAGCTGCACACTCAGCAACAGACACAGCCGCTGACCCCGCCTCCAGCTC	1653
1699	Db		GTCCAGCTGCACACTCAGCAACAGACACAGCCGCTGACCCCGCCTCCAGCTC	1758
1654	QY		TCAGACATCACTGGCTCCAGCAAGACACAGCCGGGAGGTGGAGCACTGGGGCGGAGAG	1713
1759	Db		TCAGACATCACTGGTTCCAGCAAGACACAGCCGGGAGGTGGAGCACTGGGGCGGAGAG	1818
1714	QY		CTCCAAAAGCCCCCGGAAAGGGGCTGTCTGATGAGTGAAGTACACACCTCGGC	1773
1819	Db		CTCCAAAAGCCCCCGGAAAGGGGCTGTCTGATGAGTGAAGTACACACCTCGGC	1878
1774	QY		CTGGTCAAGTGTCTGTAGCAAGTCAAGCAACCCCGGTGAAGATGTACAGTGCAGTAC	1833
1879	Db		CTGGTCAAGTGTCTGTAGCAAGTCAAGCAACCCCGGTGAAGATGTACAGTGCAGTAC	1938
1834	QY		AACTGTCTGACATGAGTACTGATTTACAGATGATCCCAGCCTCCAAAGGCCTTC	1893
1939	Db		AACTGTCTGACATGAGTACTGATTTACAGATGATCCCAGCCTCCAAAGGCCTTC	1998
1894	QY		GTGGTCAACACTGTGTTCAGGACTGGCTACGACTTGTGTGTGCTGGCCATGTGGAT	1953
1999	Db		GTGGTCAACACTGTGTTCAGGACTGGCTACGACTTGTGTGTGCTGGCCATGTGGAT	2058
1954	QY		GACACAGCAGCACTCAGGGCCACCAATCGTGGGCTGGCCAGTTCTTCCACCAAG	2013
2059	Db		GACACAGCAGCACTCAGGGCCACCAATCGTGGGCTGGCCAGTTCTTCCACCAAG	2118
2014	QY		GCTGACTCCCGCAGTCCATGACAGCCAGATTTCTGGGGCCACCATGATCCTG	2073
2119	Db		GCTGACTCCCGCAGTCCATGACAGCCAGATTTCTGGGGCCACCATGATCCTG	2178
2074	QY		GTCCATCGGGGCAATCATGTTGGCCAGCTGCTGGTCTTCATCGTCAATCTATGTTGGC	2133
2179	Db		GTCCATCGGGGCAATCATGTTGGCCAGCTGCTGGTCTTCATCGTCAATCTATGTTGGC	2238
2134	QY		TACAAGGTCTGCAACACAGGGCCCCAGTAAGATGGCAGGGCCCTGTCAGCAATGTGAC	2193
2239	Db		TACAAGGTCTGCAACACAGGGCCCCAGTAAGATGGCAGGGCCCTGTCAGCAATGTGAC	2298
2194	QY		TCCAGACCAACAGGGCCCCAGCCAGCCCTTCCAAAGCAGGCGCACAGCCGGGCCCCGGCC	2253
2299	Db		TCCAGACCAACAGGGCCCCAGCCAGCCCTTCCAAAGCAGGCGCACAGCCGGGCCCCGGCC	2358
2254	QY		CAGGGCCCCAGGTGTTGTCGCAACAGAGCTCTGACTTCCAGCTCCAGCCAGCTGGCC	2313
2359	Db		CAGGGCCCCAGGTGTTGTCGCAACAGAGCTCTGACTTCCAGCTCCAGCCAGCTGGCC	2418
2314	QY		CGGCCAGTACTCTCTCCAGCTCCCGGAGTGGGGAGGCTGGGGGCTGGGA	2373
2419	Db		CGGCCAGTACTCTCTCCAGCTCCCGGAGTGGGGAGGCTGGGGGCTGGGA	2478
2374	QY		CGGGCCCCCTGGAGGATCCCACCTCCGGCCCCCGCCCCAAGCCCTTACCGCCTG	2433
2479	Db		CGGGCCCCCTGGAGGCTCCCACCTCCGGCCCCCGCCCCAAGCCCTTACCGCCTG	2538
2434	QY		ATGGGGCCCTTCCTCCTGACACTCAAGAGTTCAGAGAAAGGAGAGTGTGGACTCC	2493

2539	Db		ATGGGGCCCTTCCTCGCTCCCTGAGCTCAAGAGCTCAGAGAAAGGAGGAGCTGTGGACTCC	2598
2494	QY		AGGACTCCACCGGGAGAGGGCTGGGACGTGGCCCCGGGGCCACCACTCGGACCCGAGAG	2553
2599	Db		AGGACTCCACCGGGAGAGGGCTGGGACGTGGCCCCGGGGCCACCACTCGGACCCGAGAG	2658
2554	QY		CAACTGTGGGGCCCTCGGGCCCGGGCCAGAGCCCTCTCCCTTGGCGTTGGAGGCG	2613
2659	Db		CAACTGTGGGGCCCTCGGGCCCGGGCCAGAGCCCTCTCCCTTGGCGTTGGAGGCG	2718
2614	QY		RAGGCCAARACGACACTCTTCCGACATGGGGGACTTGTGCTGCGGCGGGGGAGGG	2673
2719	Db		RAGGCCAARACGACACTCTTCCGACATGGGGGACTTGTGCTGCGGCGGGGGAGGG	2778
2674	QY		GTCTGTCCCGGGGGTTACAGTCTCTCCGGAAGGTCTCGAAATCTGACGAAAGCCGACG	2733
2779	Db		GTCTGTCCCGGGGGTTACAGTCTCTCCGGAAGGTCTCGAAATCTGACGAAAGCCGACG	2838
2734	QY		CTCTCTGTCAACGGCATGCTTTGGCCCTTTGAGGAGAGTGAACCTGTGGGGGGCCCCGGG	2793
2839	Db		CTCTCTGTCAACGGCATGCTTTGGCCCTTTGAGGAGAGTGAACCTGTGGGGGGCCCCGGG	2898
2794	QY		ACTTTTGGCAGCTCCGAAATGGGTGATGGAGACAGCTCTAGTGGGGGTGGGCATGCTC	2853
2899	Db		ACTTTTGGCAGCTCCGAAATGGGTGATGGAGACAGCTCTAGTGGGGGTGGGCATGCTC	2958
2854	QY		CCTTCTCTGTGCGCAGGGTGGG---AGAAGGGGAAAGAATCTACCTGGCAAGTGTGTG	2910
2959	Db		CCTTCTCTGTGCGCAGGGTGGG---AGAAGGGGAAAGAATCTACCTGGCAAGTGTGTG	3018
2911	QY		GAGTTTCCATGTGTGATTTACATCCAGGACAGTTTCTCTCTCCCTGTCATGTCCTCTG	2970
3019	Db		GAGTTTCCATGTGTGATTTACATCCAGGACAGTTTCTCTCTCCCTGTCATGTCCTCTG	3078
2971	QY		GT-----CCCCCTTACCCCGCAACACCACTACCTCCACCAACCGGGGGTGT	3025
3079	Db		GT-----CCCCCTTACCCCGCAACACCACTACCTCCACCAACCGGGGGTGT	3138
3026	QY		GCTCAGGGAATGTGGACTCGCTCAAATGCGGAGTGGCCCTGAGTGTGTGGAAGCGCA	3085
3139	Db		GCTCAGGGAATGTGGACTCGCTCAAATGCGGAGTGGCCCTGAGTGTGTGGAAGCGCA	3198
3086	QY		GACTCGCCCTTCTAATCAAAATGATGACCAAGGAGGCTTGTGATGCTTATG	3144
3199	Db		GACTCGCCCTTCTAATCAAAATGATGACCAAGGAGGCTTGTGATGCTTATG	3257

RESULT 3
AL591063/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Human DNA sequence from clone Rpl1-535K1 on chromosome 6, complete sequence.
AL591063
AL591063.6 GI:15020551
HTG.
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 100426)
Direct Submission
Submitted (19-JUL-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
Request: clonerequest@sanger.ac.uk
On Jul 25, 2001 this sequence version replaced gi:14787623.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows otherwise noted: all

regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em.; EMBL; SW.; SWISSPROT; Tr.; TREMBL; Wp.; WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/projects/C_elegans/wormpep This sequence was generated from part of bacterial clone contigs of human chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/Chr6> RP11-535K1 is from the library RP11-11.2 constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm> VECTOR: pBACE3.6

IMPORTANT: This sequence is not the entire insert of clone RP11-535K1 It may be shorter because we sequence overlapping sections only once, except for a 100 base overlap. The true left end of clone RP11-121P10 is at 98427 in this sequence. The true right end of clone RP11-552E20 is at 2000 in this sequence.

FEATURES

source Location/Qualifiers

1..100426
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="6"
 /clone="RP11-535K1"
 /clone_lib="RP11-11.2"

1..1026
 /note="L1M3 repeat: matches 4727. .5917 of consensus"
 repeat_region

1030..1221
 /note="L1 repeat: matches 4267. .4460 of consensus"
 repeat_region

1222..1296
 /note="Alu repeat: matches 1. .50 of consensus"
 repeat_region

1297..3135
 /note="L1 repeat: matches 2379. .4267 of consensus"
 repeat_region

3134..3536
 /note="L1M1 repeat: matches 5186. .5588 of consensus"
 repeat_region

3537..3675
 /note="FLAM A repeat: matches 1. .142 of consensus"
 repeat_region

3677..4233
 /note="L1M1 repeat: matches 5582. .6164 of consensus"
 repeat_region

4523..5546
 /note="L1M4 repeat: matches 2429. .3489 of consensus"
 repeat_region

5510..5956
 /note="L1M5 repeat: matches 4780. .5243 of consensus"
 repeat_region

5997..6289
 /note="MT1A1 repeat: matches 102. .365 of consensus"
 repeat_region

6303..7075
 /note="L1M5 repeat: matches 5264. .6022 of consensus"
 repeat_region

7078..7139
 /note="MSTA repeat: matches 365. .426 of consensus"
 repeat_region

7147..7205
 /note="Alu repeat: matches 244. .302 of consensus"
 repeat_region

7201..7553
 /note="MSTA repeat: matches 1. .366 of consensus"
 repeat_region

7555..7710
 /note="L1M6 repeat: matches 6020. .6172 of consensus"
 repeat_region

9233..9423
 /note="MIR repeat: matches 22. .219 of consensus"
 repeat_region

9441..9522
 /note="41 copies 2 mer tt 63% conserved"
 repeat_region

11776..11955
 /note="MIR repeat: matches 59. .235 of consensus"
 repeat_region

15525..15661
 /note="L2 repeat: matches 2594. .2744 of consensus"
 repeat_region

15837..16034
 /note="MIR repeat: matches 60. .262 of consensus"
 repeat_region

18523..18696
 /note="MIR repeat: matches 78. .256 of consensus"
 repeat_region

18748..19107
 /note="MRR33 repeat: matches 3. .324 of consensus"
 repeat_region

20177..20290
 /note="3 copies 38 mer 76% conserved"
 repeat_region

20885..20965
 /note="MIR repeat: matches 35. .115 of consensus"
 repeat_region

22278..22945
 /note="CpG island"
 /evidence=not_experimental
 repeat_region

23265..23372
 /note="18 copies 6 mer agagag 64% conserved"
 repeat_region

23279..23372
 /note="47 copies 2 mer ag 67% conserved"
 repeat_region

23488..23613
 /note="63 copies 2 mer ga 58% conserved"
 repeat_region

23498..23611
 /note="19 copies 6 mer gacaga 63% conserved"
 repeat_region

23532..23603
 /note="4 copies 18 mer 88% conserved"
 repeat_region

23533..23604
 /note="2 copies 36 mer 90% conserved"
 repeat_region

24897..25274
 /note="THL1B repeat: matches 1. .364 of consensus"
 repeat_region

25405..25583
 /note="MIR repeat: matches 10. .180 of consensus"
 repeat_region

25906..26082
 /note="MIR repeat: matches 14. .218 of consensus"
 repeat_region

26339..26599
 /note="MIR repeat: matches 2. .262 of consensus"
 repeat_region

26933..27010
 /note="13 copies 6 mer tgtgtg 92% conserved"
 repeat_region

27107..27225
 /note="MIR repeat: matches 120. .262 of consensus"
 repeat_region

27481..27801
 /note="L2 repeat: matches 1904. .2246 of consensus"
 repeat_region

28341..28546
 /note="MER20 repeat: matches 9. .218 of consensus"
 repeat_region

28712..28786
 /note="LTR33 repeat: matches 10. .87 of consensus"
 repeat_region

29265..29354
 /note="LTR33 repeat: matches 110. .198 of consensus"
 repeat_region

29670..29983
 /note="AluX repeat: matches 1. .312 of consensus"
 repeat_region

30423..30512
 /note="L2 repeat: matches 2602. .2695 of consensus"
 repeat_region

30662..30966
 /note="AluX repeat: matches 1. .305 of consensus"
 repeat_region

31122..31173
 /note="26 copies 2 mer tg 71% conserved"
 repeat_region

31489..31573
 /note="MTLJ repeat: matches 103. .190 of consensus"
 repeat_region

32390..32444
 /note="L2 repeat: matches 2651. .2705 of consensus"
 repeat_region

32405..32618
 /note="MIR repeat: matches 3. .249 of consensus"
 repeat_region

32729..32977
 /note="MIR repeat: matches 14. .262 of consensus"
 repeat_region

33132..33310
 /note="L2 repeat: matches 2254. .2449 of consensus"
 repeat_region

33922..34526
 /note="L2 repeat: matches 2059. .2672 of consensus"
 repeat_region

34553..34647
 /note="L2 repeat: matches 2658. .2749 of consensus"
 repeat_region

34662..34869
 /note="MER3 repeat: matches 1. .209 of consensus"
 repeat_region

35504..35771
 /note="MIR repeat: matches 2. .257 of consensus"
 repeat_region

35720..35782
 /note="L2 repeat: matches 2688. .2750 of consensus"
 repeat_region

36322..36529
 /note="MIR repeat: matches 22. .255 of consensus"
 repeat_region

37340..37724

```

/note="MIR repeat: matches 4. .188 of consensus"
repeat_region 38100..38197
/note="U6 repeat: matches 2. .103 of consensus"
repeat_region 38228..38496
/note="L1MB3 repeat: matches 5903. .6181 of consensus"
repeat_region 38815..38998
/note="MIR repeat: matches 8. .228 of consensus"
repeat_region 39066..39383
/note="MSTD repeat: matches 1. .394 of consensus"
repeat_region 39778..39889
/note="MIR repeat: matches 53. .165 of consensus"
repeat_region 40464..40529
/note="MIR repeat: matches 51. .120 of consensus"
repeat_region 40657..40862
/note="MIR repeat: matches 20. .249 of consensus"
repeat_region 41569..41640
/note="36 copies 2 mer gt 66% conserved"
repeat_region 41737..42201
/note="L2 repeat: matches 2203. .2701 of consensus"
repeat_region 42225..42370
/note="MIR repeat: matches 116. .262 of consensus"
repeat_region 42733..42853
/note="MIR repeat: matches 13. .128 of consensus"
repeat_region 42907..43213
/note="L2 repeat: matches 2452. .2750 of consensus"
repeat_region 44325..44471
/note="MIR repeat: matches 31. .184 of consensus"
repeat_region 44544..44770
/note="MIR repeat: matches 27. .262 of consensus"
repeat_region 45462..45770

```

```

Query Match 45.3%; Score 1424.8; DB 9; Length 100426;
Best Local Similarity 97.2%; Pred. No. 3,9e-237;
Matches 1450; Conservative 0; Mismatches 42; Indels 0; Gaps 0;
QY 420 ACCGGGGTAGATGCTGCCGCCAGGGCTGAGTGACCAGACCAATGGAGACCCCTGCT 479
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
63405 ACCCTGAGCTTGTGTTCCCTTCCAGGCGGTGAGTGACCAGACCAATGGAGACCCCTGCT 63346
QY 480 TGGTGCCCTGACTGAGTTTGGATGGCTTGGATGGCTGCGGCTGAGCGCTGCCAAGTACTG 539
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
63345 TGGTGCCCTGACTGAGTTTGGATGGCTTGGATGGCTGCGGCTGAGCGCTGCCAAGTACTG 63286

```

```

QY 540 TGTCTGCCAGATCTGTGAGTCACTGGGGACCCTGTGCCCTCCAGGGGCTGCTCT 599
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
63285 TGTCTGCCAGATCTGTGAGTCACTGGGGACCCTGTGCCCTCCAGGGGCTGCTCT 63226
QY 600 TGTATCCCTGTATTTACCGGGACAGTGGAGCTGCCTGGGGGCAACTTCAAT 659
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
63225 TGTATCCCTGTATTTACCGGGACAGTGGAGCTGCCTGGGGGCAACTTCAAT 63166
QY 660 CCACATCAGCCGACGACTTTCGCAACATGACGGGGCTGTTGACCTGACCTGTCACG 719
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
63165 CCACATCAGCCGACGACTTTCGCAACATGACGGGGCTGTTGACCTGACCTGTCACG 63106
QY 720 GAACACATCAGCCATCCAGCCCTTTCTTTCGTGGACCTGACGAGCTCCGCTCCCT 779
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
63105 GAACACATCAGCCATCCAGCCCTTTCTTTCGTGGACCTGACGAGCTCCGCTCCCT 63046
QY 780 GCATTTGACGAAATCGGCTGCAAGCTTGGGAGGACACCCTCCGGGGCTGTTGCA 839
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
63045 GCATTTGACGAAATCGGCTGCAAGCTTGGGAGGACACCCTCCGGGGCTGTTGCA 62986
QY 840 CCTGGACACCTTATCGTGAACAACAACAGCTGGGGCCATCGCATGGCTTTTCA 899
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
62985 CCTGGACACCTTATCGTGAACAACAACAGCTGGGGCCATCGCATGGCTTTTCA 62926
QY 900 GGACTTCTGCTGCATTTGGAGATCTGGACCTCTCTTACAACAACTCCATGGCTGC 959
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
62925 GGACTTCTGCTGCATTTGGAGATCTGGACCTCTCTTACAACAACTCCATGGCTGC 62866
QY 960 GTGGACTCCGTCGACGATGTTCAACTCCACAGCTGAGCTTGGTCAACAACTTGT 1019
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

```

Db 62865 GTGGACTCCGTCGACGCAATGGTCAAACTCCACCAGCTGACCTGGACCACAACTGCT 62806
QY 1020 GGATCACATCGCCGAGGGCACCCTTTGACAGACTGAGAACTGGCCCGCTGGATCTCAC 1079
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
62805 GGATCACATCGCCGAGGGCACCCTTTGACAGACTGAGAACTGGCCCGCTGGATCTCAC 62746
QY 1080 CTCGAATCGGCTGCAAGAGCTCCCTGATCCCACTTTTGCCCGCTCCAGGCTTCGGC 1139
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
62745 CTCGAATCGGCTGCAAGAGCTCCCTGATCCCACTTTTGCCCGCTCCAGGCTTCGGC 62686
QY 1140 TTTGACAGCCACACCCTTTTGGCCCCACCCTTTCCTTTAGTTTGGGGTAACCCACTTCA 1199
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
62685 TTTGACAGCCACACCCTTTTGGCCCCACCCTTTCCTTTAGTTTGGGGTAACCCACTTCA 62626
QY 1200 CTCGAATGTTGAGCTTCTTGCTGGAGGGCTGCGAGGGGCGACGATGACTGGAACTCG 1259
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
62625 CTCGAATGTTGAGCTTCTTGCTGGAGGGCTGCGAGGGGCGACGATGACTGGAACTCG 62566
QY 1260 TGGCTCCCAGGGGCTCAAGGGTGCCTACTTTGCGCATGTGCGTGGAGAGTGTGT 1319
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
62565 TGGCTCCCAGGGGCTCAAGGGTGCCTACTTTGCGCATGTGCGTGGAGAGTGTGT 62506
QY 1320 GTGCGAGCCGCTCCTATCACCCAGCACACACAAGTGTGGTTCGAGGCCCCAGC 1379
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
62505 GTGCGAGCCGCTCCTATCACCCAGCACACACAAGTGTGGTTCGAGGCCCCAGC 62446
QY 1380 GGCCACACTCAAGTGCAAAAGCCATTTGGGACCCCGCCCTTATCCACTGGGTAGCCCC 1439
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
62445 GGCCACACTCAAGTGCAAAAGCCATTTGGGACCCCGCCCTTATCCACTGGGTAGCCCC 62386
QY 1440 CGATGACCGCTGGTAGGAACTCTCAAGGACCGCTGTGTATGACAAATGGAACCCCTGGA 1499
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
62385 CGATGACCGCTGGTAGGAACTCTCAAGGACCGCTGTGTATGACAAATGGAACCCCTGGA 62326
QY 1500 CATCTCATCACACATCTCAGGACAGTGGCTTCCCTGCTGCATGCTGCCAATGCTGC 1559
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
62325 CATCTCATCACACATCTCAGGACAGTGGCTTCCCTGCTGCATGCTGCCAATGCTGC 62266
QY 1560 CCGAGAGCCACGGCCATGGTGGAGTCTCAATGCTCCAGCTGCCACACTCAGCAACAG 1619
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
62265 CCGAGAGCCACGGCCATGGTGGAGTCTCAATGCTCCAGCTGCCACACTCAGCAACAG 62206
QY 1620 CACCAGCCGACTGACCCCGCCCAAGTCCCGCTCCTCAGACACTCACTGGCTCCAGCAAGAC 1679
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
62205 CACCAGCCGACTGACCCCGCCCAAGTCCCGCTCCTCAGACACTCACTGGCTCCAGCAAGAC 62146
QY 1680 CAGCCGGGAGTGGAGGAGTGGGGCGGAGAGCTCCAAAGCCCCCGGAAAGGGC 1739
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
62145 CAGCCGGGAGTGGAGGAGTGGGGCGGAGAGCTCCAAAGCCCCCGGAAAGGGC 62086
QY 1740 TGTGCTTGTGCTGAAGTGCACCACCCTCCCGCCCTGCTGCAAGTGTGCTGTCAGCAAGTC 1799
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
62085 TGTGCTTGTGCTGAAGTGCACCACCCTCCCGCCCTGCTGCAAGTGTGCTGTCAGCAAGTC 62026
QY 1800 AGCACCCTGGTGGAAGTGTACCAGCTGCAGTACAATGCTCTGACGATGAGGTACTGAT 1859
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
62025 AGCACCCTGGTGGAAGTGTACCAGCTGCAGTACAATGCTCTGACGATGAGGTACTGAT 61966
QY 1860 TTACAGGATGATCCCAGCTCCAAACAGCCCTTCGTTGTCACAACTCTGGT 1911
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
61965 TTACAGGATGATCCCAGCTCCAAACAGCCCTTCGTTGTCACAACTCTGGT 61914

```

```

RESULT 4
AC011679 179820 bp DNA linear HTG 26-MAR-2000
LOCUS Homo sapiens clone RP11-14G23, WORKING DRAFT SEQUENCE, 16 unordered
DEFINITION pieces.
ACCESSION AC011679
VERSION AC011679.3 GI:7329391
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens.
ORGANISM Homo sapiens

```

```

* 49447 49546: gap of 100 bp
* 49547 60664: contig of 11118 bp in length
* 60665 60764: gap of 100 bp
* 60765 74690: contig of 13926 bp in length
* 74691 74790: gap of 100 bp
* 74791 89780: contig of 14990 bp in length
* 89781 89880: gap of 100 bp
* 89881 109807: contig of 19927 bp in length
* 109808 109907: gap of 100 bp
* 109908 128003: contig of 18096 bp in length
* 128004 128103: gap of 100 bp
* 128104 148838: contig of 20735 bp in length
* 148839 148938: gap of 100 bp
* 148939 179820: contig of 30882 bp in length.

```

FEATURES

```

source
1..179820
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="RP11-14G23"
/clone_lib="RPCI-11 Human Male BAC"
1..1960
/feature="assembly_fragment"
2061..3217
/feature="assembly_fragment"
clone_end:SP6
vector_side:left
3318..6117
/feature="assembly_fragment"
6218..8883
/feature="assembly_fragment"
8984..12533
/feature="assembly_fragment"
12634..17388
/feature="assembly_fragment"
17489..27094
/feature="assembly_fragment"
27195..38372
/feature="assembly_fragment"
38473..49446
/feature="assembly_fragment"
49547..60664
/feature="assembly_fragment"
60765..74690
/feature="assembly_fragment"
74791..89780
/feature="assembly_fragment"
89881..109807
/feature="assembly_fragment"
109908..128003
/feature="assembly_fragment"
clone_end:I7
vector_side:right
128104..148838
/feature="assembly_fragment"
148939..179820
/feature="assembly_fragment"
46444 a 43444 c 42485 g 45943 t 1504 others

```

```

BASE COUNT 46444 a 43444 c 42485 g 45943 t 1504 others
ORIGIN

```

```

Query Match 40.8%; Score 1282; DB 2; Length 179820;
Best Local Similarity 100.0%; Pred. No. 2e-212;
Matches 1282; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 1863 CAGGATATCCAGCCCTCCACCAAGCCCTTCGGTCAACAACTGGTGTCCAGGACTGG 1922
Db 112240 CAGGATATCCAGCCCTCCACCAAGCCCTTCGGTGTCAACAACTGGTGTCCAGGACTGG 112299
Oy 1923 CTAGCACTTGTGTGTCTGGCCATGTGGATGACACAGCCACCACTACGGCCACCAA 1982
Db 112300 CTAGCACTTGTGTGTCTGGCCATGTGGATGACACAGCCACCACTACGGCCACCAA 112359
Oy 1983 CATCGTGGCTGCCCCAGTTCITCACCAAGGTCATTCCAGGTCACCCGAGTCCAGTCA 2042

```

```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 179820)
Birren,B., Linton,L., Nusbaum,C. and Lander,E.
Homo sapiens, clone RP11-14G23
Unpublished
2. (bases 1 to 179820)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
Baldwin,J., Barna,N., Beckerly,R., Boguslavkiy,L., Boukhgalter,B.,
Brown,A., Castle,A., Colangelo,M., Collins,S., Collymore,A.,
Cooke,P., DeArelano,K., Dewar,K., Domino,M., Donelan,L., Doyle,M.,
Ferreira,P., FitzHugh,W., Forrest,C., Funke,R., Gage,D.,
Galagan,J., Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
Lehoczky,J., Lieu,C., Locke,K., Macdonald,P., Marquis,N.,
McEwan,P., McGurk,A., McKernan,K., McLaughlin,J., Meldrum,J.,
Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P.,
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Teeffay,S., Tirrell,A., Vassiliev,H., Vo,A., Wheeler,J., Wu,X.,
Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.
Direct Submission
Submitted (10-OCT-1999) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Mar 26, 2000 this sequence version replaced gi:6249709.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www.seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L3400
Center clone name: 14.G.23
----- Summary Statistics
Sequencing vector: M13; M77815; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 166032 bases at least Q40
Consensus quality: 172101 bases at least Q30
Consensus quality: 174784 bases at least Q20
Insert size: 182000; agarose-fp
Insert size: 178320; sum-of-contigs
Quality coverage: 4.1 in Q20 bases; agarose-fp
Quality coverage: 4.2 in Q20 bases; sum-of-contigs

```

COMMENT

```

* NOTE: This is a 'working draft' sequence. It currently
* consists of 16 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 1960: contig of 1960 bp in length
* 1961 2060: gap of 100 bp
* 2061 3217: contig of 1157 bp in length
* 3218 3317: gap of 100 bp
* 3318 6117: contig of 2800 bp in length
* 6118 6217: gap of 100 bp
* 6218 8883: contig of 2666 bp in length
* 8884 8983: gap of 100 bp
* 8984 12533: contig of 3550 bp in length
* 12534 12633: gap of 100 bp
* 12634 17388: contig of 4755 bp in length
* 17389 17488: gap of 100 bp
* 17489 27094: contig of 9606 bp in length
* 27095 27194: gap of 100 bp
* 27195 38372: contig of 11178 bp in length
* 38373 38472: gap of 100 bp
* 38473 49446: contig of 10974 bp in length

```

```

REFERENCE
AUTHORS
TITLE
JOURNAL
AUTHORS

```

```

TITLE
JOURNAL

```

COMMENT

Db 112360 CATCTGGGCTGGCCCGATCTTTCACCAAGGCTACTACCGCAGTGCAGTCCATGCA 112419
 QY 2043 CAGCCAGATTTGGCGGCACCATGATCTCTGTCATCGGGGGCATCATCTGGCCACGCT 2102
 Db 112420 CAGCCAGATTTGGCGGCACCATGATCTCTGTCATCGGGGGCATCATCTGGCCACGCT 112479
 QY 2103 GCTGTTTCATCTCATCTCTCATGTTGGCTTACAAGTTTGCACCAACAGGCCCCCAG 2162
 Db 112480 GCTGTTTCATCTCATCTCTCATGTTGGCTTACAAGTTTGCACCAACAGGCCCCCAG 112539
 QY 2163 CAAGATGGCAGGGCCCTGAGCAATGTGTACTCGCAGCAACAAGGGCCCGCCAGCCCGCC 2222
 Db 112540 CAAGATGGCAGGGCCCTGAGCAATGTGTACTCGCAGCAACAAGGGCCCGCCAGCCCGCC 112599
 QY 2223 TCAAAGCAGCCACCAAGCCCGGGCCCGCCGACGAGCCCGCCGAAAGGTTGGTGGCGCAA 2282
 Db 112600 TCAAAGCAGCCACCAAGCCCGGGCCCGCCGACGAGCCCGCCGAAAGGTTGGTGGCGCAA 112659
 QY 2283 CGAGTCTTGGACTTACCGCCAGCCCTGGCCCGCCGACGAGTCTCTTCTTCCAGCTC 2342
 Db 112660 CGAGTCTTGGACTTACCGCCAGCCCTGGCCCGCCGACGAGTCTCTTCTTCCAGCTC 112719
 QY 2343 CTTGGCAGTGGGGAGCTGGGGGCTGGAGCGGGCCCTGGAGGATCCACCCCTCCG 2402
 Db 112720 CTTGGCAGTGGGGAGCTGGGGGCTGGAGCGGGCCCTGGAGGATCCACCCCTCCG 112779
 QY 2403 CCCGCCCCAAAGCCAGCCCTTTGACCCCTGTATGGGGGCTTGGCCCTCCCTGACCTCAA 2462
 Db 112780 CCCGCCCCAAAGCCAGCCCTTTGACCCCTGTATGGGGGCTTGGCCCTCCCTGACCTCAA 112839
 QY 2463 GAGTCAGAAAAGAGGAGCTGCTGACTCCAGGACTCAGCCGGAGAGGGGCTGGGAC 2522
 Db 112840 GAGTCAGAAAAGAGGAGCTGCTGACTCCAGGACTCAGCCGGAGAGGGGCTGGGAC 112899
 QY 2523 GTCGGCCGGCCGACCACTCGAGCCAGAGCAGCTGTGGGGCCCGCCGCGCCGGCC 2582
 Db 112900 GTCGGCCGGCCGACCACTCGAGCCAGAGCAGCTGTGGGGCCCGCCGCGCCGGCC 112959
 QY 2583 CAGGAGCTCTCCCTTCCGTTGGAGGCAAGCCAAAGCAGCAGCTCTTCCGAT 2642
 Db 112960 CAGGAGCTCTCCCTTCCGTTGGAGGCAAGCCAAAGCAGCAGCTCTTCCGAT 113019
 QY 2643 GGGGACTTGTCTGTCGGCCGGGAGGGTGTGTCGGGGGCTACAGTCTCTCTCG 2702
 Db 113020 GGGGACTTGTCTGTCGGCCGGGAGGGTGTGTCGGGGGCTACAGTCTCTCTCG 113079
 QY 2703 GAAGTCTGACATCTGACGAAGCAGCCCTCTGTCAAGGGATCTCTTGGCCCT 2762
 Db 113080 GAAGTCTGACATCTGACGAAGCAGCCCTCTGTCAAGGGATCTCTTGGCCCT 113139
 QY 2763 TGAGGAGATGACTTGGGGCCCGGGGACTTTTGGCAGCTCCGAAATGGGTGATGGA 2822
 Db 113140 TGAGGAGATGACTTGGGGCCCGGGGACTTTTGGCAGCTCCGAAATGGGTGATGGA 113199
 QY 2823 GAGCAGGCTTAGTGGGGTGGGATGTCTTCTTCTTCTGTCGCGAGGTTGGGAAAGGG 2882
 Db 113200 GAGCAGGCTTAGTGGGGTGGGATGTCTTCTTCTTCTGTCGCGAGGTTGGGAAAGGG 113259
 QY 2883 GAAAGAATCTCACTGCAAGTGTGGAGTTCATGGTGTTCATCCAGGGAC 2942
 Db 113260 GAAAGAATCTCACTGCAAGTGTGGAGTTCATGGTGTTCATCCAGGGAC 113319
 QY 2943 AGTTTGGTCTCCCTGTCAATGGCCCTGTGTCCCGCCCTACCCCGCAACCCACATCACCC 3002
 Db 113320 AGTTTGGTCTCCCTGTCAATGGCCCTGTGTCCCGCCCTACCCCGCAACCCACATCACCC 113379
 QY 3003 TCCCCACCCCGCCGGGTTGTCTCAGGAAATGAGACTCGCTCAAAATGCCGGACTGA 3062
 Db 113380 TCCCCACCCCGCCGGGTTGTCTCAGGAAATGAGACTCGCTCAAAATGCCGGACTGA 113439
 QY 3063 GCCTCAGTCTTGGAAAAGCCGAGACTCCGCTTCTTAATCAAAATGAGCTCAAGC 3122
 Db 113440 GCCTCAGTCTTGGAAAAGCCGAGACTCCGCTTCTTAATCAAAATGAGCTCAAGC 113499

QY 3123 AAGCGCTTTGGATTGCTTATG 3144
 Db 113500 AAGCGCTTTGGATTGCTTATG 113521

RESULT 5
 AC101792
 LOCUS
 DEFINITION
 AC101792 159443 bp DNA linear HTG 21-AUG-2002
 Mus musculus clone RP24-176G22, WORKING DRAFT SEQUENCE, 20
 unordered pieces.
 AC101792
 HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
 house mouse.
 Mus musculus
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 Birren,B., Nusbaum,C. and Lander,B.
 Mus musculus, clone RP24-176G22
 Unpublished
 2 (bases 1 to 159443)
 Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
 Anderson,S., Barna,N., Bastien,V., Boguslavkiy,L., Boukhgalter,B.,
 Brown,A., Camarata,J., Campolano,A., Chang,J., Chazaro,B.,
 Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A.,
 Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
 Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
 Ginde,S., Gurd,S., Goyette,M., Graham,L., Grand-Pierre,N.,
 Hagos,B., Heathford,A., Horton,L., Hulme,W., Iliiev,I., Johnson,R.,
 Jones,C., Kamat,A., Karatas,A., Kells,C., LaRoque,K.,
 Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Liu,G.,
 MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
 McCarthy,M., McEwan,P., McKernan,K., MCPheeters,R., Meldrim,J.,
 Mensu,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C.,
 Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D.,
 Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
 Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
 Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R.,
 Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
 Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
 Topham,K., Travers,M., Travis,N., Triggilio,J., Vassiliev,H.,
 Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
 Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

Submitted (23-NOV-2001) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 3 (bases 1 to 159443)
 Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
 Barna,N., Bastien,V., Bloom,T., Boguslavkiy,L., Boukhgalter,B.,
 Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,
 Cooke,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
 Faro,S., Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J.,
 Gardyna,S., Gurd,S., Graham,L., Grand-Pierre,N., Hagos,B.,
 Horton,L., Hulme,W., Iliiev,I., Johnson,R., Jones,C., Kamat,A.,
 Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K.,
 Liu,G., MacLean,C., Macdonald,P., Major,J., Matthews,C.,
 McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,
 Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H.,
 O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K.,
 Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P.,
 Roman,J., Roy,A., Schauer,S., Schupback,R., Seaman,S., Severy,P.,
 Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J.,
 Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H.,
 Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J.,
 Zembek,L., Zimmer,A. and Zody,M.

TITLE
 JOURNAL
 REFERENCE
 AUTHORS

Direct Submission
 Submitted (23-NOV-2001) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 3 (bases 1 to 159443)
 Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
 Barna,N., Bastien,V., Bloom,T., Boguslavkiy,L., Boukhgalter,B.,
 Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,
 Cooke,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
 Faro,S., Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J.,
 Gardyna,S., Gurd,S., Graham,L., Grand-Pierre,N., Hagos,B.,
 Horton,L., Hulme,W., Iliiev,I., Johnson,R., Jones,C., Kamat,A.,
 Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K.,
 Liu,G., MacLean,C., Macdonald,P., Major,J., Matthews,C.,
 McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,
 Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H.,
 O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K.,
 Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P.,
 Roman,J., Roy,A., Schauer,S., Schupback,R., Seaman,S., Severy,P.,
 Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J.,
 Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H.,
 Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J.,
 Zembek,L., Zimmer,A. and Zody,M.

TITLE
 JOURNAL
 COMMENT

Direct Submission
 Submitted (21-AUG-2002) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 On Aug 21, 2002 this sequence version replaced gi:17060567.
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)

/clone_lib="RPCI-24 Male Mouse BAC"

1..1129
/note="assembly_fragment"
1230..2261
/note="assembly_fragment"
2362..3703
/note="assembly_fragment"
3804..5915
/note="assembly_fragment"
6016..8269
/note="assembly_fragment"
8370..9416
/note="assembly_fragment"
9517..12017
/note="assembly_fragment"
12118..13953
/note="assembly_fragment"
14054..16946
/note="assembly_fragment"
17047..20068
/note="assembly_fragment"
20169..22610
/note="assembly_fragment"
22711..25193
/note="assembly_fragment"
25294..28313
/note="assembly_fragment"
28414..31724
/note="assembly_fragment"
31825..33932
/note="assembly_fragment"
40033..50334
/note="assembly_fragment"
50435..66250
/note="assembly_fragment"
66351..78871
/note="assembly_fragment"
78972..118433
/note="assembly_fragment"
118534..159443
/note="assembly_fragment"
1902 others

BASE COUNT 41724 a 36746 c 35340 g 43731 t
ORIGIN
Query Match 37.1%; Score 1166.6; DB 2; Length 159443;
Best Local Similarity 86.6%; Pred. No. 2e-192;
Matches 1286; Conservative 0; Mismatches 199; Indels 0; Gaps 0;

QY 426 GGTAGATCGTCCCTCGCCAGCGCTGAGTGACAGACCCATGGAGACCCCTGCTGTGGTGG 485
Db 90878 GGTGCTCTTTCCAGGTGACAGAGCTGAGTGACAGACCCATGGAGACTGCTGTGGTGG 90937
QY 486 CTTGTAGCGTTGGCATGGCGTTTCCCGTGGTGCAGCGCTGCCCAAGTACTGTGTCTG 545
Db 90938 GCTGTGGTGGTGGCATGGCGTTTGTGTGGTGCATGCTGCCCAAGTACTGTGTCTG 90997
QY 546 CCAAGATGCTGTGAGTCACTGCGGACCCCTGTGCCCTTCAAGGGGTGCTCTTTGTACC 605
Db 90998 CCAAGATGCTGTGAGTCACTGCGGACCCCTGTGCCCTTCAAGGGGTGCTCTTTGTACC 91057
QY 606 CCCTGATATTGACCGCGGACAGTGGAGCTGCGCTGGCGGCACTTCAATCCACAT 665
Db 91058 CCCTGATATTGACCGCGGACAGTGGAGCTGCGCTGGCGGCACTTCAATCCACAT 91117
QY 666 CAGCCCGCAGACTTTGCCAACATGACGGGGCTGGTGGACCTGACCCCTGTCAGGAACAC 725
Db 91118 TGGCGCCAGACTTTGCCAACATGACGGGGCTGGTGGACCTGACCCCTGTCAGGAACAC 91177
QY 726 CATCAGCCACATCCAGCCCTTTTCCCTTTTGGACCTCGAGGCTCGGCTCCCTGCACT 785
Db 91178 CATCAGCCACATCCAGCCCTTTTCCCTTTTGGACCTCGAGGCTCGGCTCCCTGCACT 91237
QY 786 TGACAGCAATCGGCTGCCAAGCCCTTGGGAGGACACCCCTCCGGGGCTGTGTCACCTGCA 845

http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L17407
Center clone name: 176 G 22
----- Summary Statistics

Sequencing vector: Plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 152592 bases at least Q40
Consensus quality: 155716 bases at least Q30
Consensus quality: 156921 bases at least Q20
Insert size: 157000; agarose-fp
Insert size: 157543; sum-of-contigs
Quality coverage: 7.8 in Q20 bases; agarose-fp
Quality coverage: 7.8 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 20 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 1129: contig of 1129 bp in length
1130 1229: gap of 100 bp
1230 2261: contig of 1032 bp in length
2262 2361: gap of 100 bp
2362 2361: gap of 100 bp
2362 3703: contig of 1342 bp in length
3704 3803: gap of 100 bp
3804 5915: contig of 2112 bp in length
5916 6015: gap of 100 bp
6016 8269: contig of 2254 bp in length
8270 8369: gap of 100 bp
8370 9416: contig of 1047 bp in length
9417 9516: gap of 100 bp
9517 12017: contig of 2501 bp in length
12018 12117: gap of 100 bp
12118 13953: contig of 1836 bp in length
13954 14053: gap of 100 bp
14054 16946: contig of 2893 bp in length
16947 17046: gap of 100 bp
17047 20068: contig of 3022 bp in length
20069 20168: gap of 100 bp
20169 22610: contig of 2442 bp in length
22611 22710: gap of 100 bp
22711 25193: contig of 2483 bp in length
25194 25293: gap of 100 bp
25294 28313: contig of 3020 bp in length
28314 28413: gap of 100 bp
28414 31724: contig of 3311 bp in length
31725 31824: gap of 100 bp
31825 39932: contig of 8108 bp in length
39933 40032: gap of 100 bp
40033 50334: contig of 10302 bp in length
50335 50434: gap of 100 bp
50435 66250: contig of 15816 bp in length
66251 66350: gap of 100 bp
66351 78871: contig of 12521 bp in length
78872 78971: gap of 100 bp
78972 118433: contig of 39462 bp in length
118434 118533: gap of 100 bp
118534 159443: contig of 40910 bp in length.
Location/Qualifiers
1 159443
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="RP24-176G22"

FEATURES
source

Db 91238 TGACAGCAACCGGCTACCTAGCCCTTGGGAGGACACACTCCGGGGTCTGGTCAACCTGCA 91297
 QY 846 GCACCTTATCGTGAACAACAACAGCTGGCGGCATCGCAGATGAGGCTTTTGAGACTT 905
 Db 91298 ACACCTTATGTGAACAATAACAGCTGGGTGGCATCGTGATGATGCCTTTGAAGACTT 91357
 QY 906 CTTGCTGACATTTGGAGATCTGGACCTCTCTTACAAACAACCTCCATCGGCTGGCGTGGGA 965
 Db 91358 CTTGCTGACTTTAGAAAGACTGGACCTATCTTACAAACAACCTCCAGGACTTCCCTGGGA 91417
 QY 966 CTTGCTGACGATGCTGGAGATCTGGACCTCTCTTACAAACAACCTCCATCGGCTGGCGTGGGA 1025
 Db 91418 TTCCGTPACGGGATGTTCAACCTCTTACAGCTGATCTGGACCAACAATCTGCTGGACCA 91477
 QY 1026 CATCGCGGAGGACCTTTGCAGACCTGCAGAAACTGGCCCGCTGGATCTCACCTCCAA 1085
 Db 91478 CATGCTGAGGACCTTTGCAGACCTGCAGAAACTGGCCCGCTGGACCTCACATCCAA 91537
 QY 1086 TGGCTGCAAGAGCTCCCGCTGATCCCAATCTTTGCGCCGCTCCAGGCTTGGCTTTGAC 1145
 Db 91538 TCGGCTGCAGAAACTTCCCGGACCCCTCTTTGCGCCGCTCCAGGCTTCCCTGCTAAC 91597
 QY 1146 AGCCACACCTTTGCCCCACCTTGTCTTCTTTAGTTTGGGTTAACCCACTTCACTGCAA 1205
 Db 91598 TGGCACGCTTTGCCCCACCTTGTCTTCTTTAGTTTGGGTTAACCCACTTCACTGCAA 91657
 QY 1206 TTTGAGCTTCTTGGCTGGAGGCTCGAGCGGACGATGACCTGGAACCTTGGGCTC 1265
 Db 91658 TTTGAGCTTCTTGGCTGGAGGCTCGAGCGGACGATGACCTGGAACCTTGGGCTC 91717
 QY 1266 CCAGGGGCTCAAGGCTCGTACTTTGGCATGTGGCATGTGGAGGAGTTTGTGTGGA 1325
 Db 91718 CCCCCGAAGTCTCAAGGCTCGTACTTTGGCATTTGGCATTTGGAGGAGTTTGTGTGGA 91777
 QY 1326 GCGCTCTCATCACCCAGCACACACAAGTTGCTGTTTGGAGGCGAGGCGGCCAC 1385
 Db 91778 GCGCTCTCATCACCCAGCACACACAAGTTGCTGTTTGGAGGCGAGGCGGCCAC 91837
 QY 1386 ACTCAAGTCAAAGGCTTGGGACCCAGCCCTTATCCACTGGGTAGCCCCCATGA 1445
 Db 91838 TCTTAAGTCAAGGCTTGGGACCCAGCCCTTATCCACTGGGTAGCCCCCATGA 91897
 QY 1446 CCGCTTGGTAGGAACTCTCAAGGACCGCTGTTATGACAATGGCACCTGGACATCTT 1505
 Db 91898 CCGCTTGGTAGGAACTCTCAAGGACCGCTGTTATGACAATGGCACCTGGACATCTT 91957
 QY 1506 CATCACCATCTCAGACAGTGGTCTTCACTGATCTGCAATGCTGCGGGAGA 1565
 Db 91958 AATTACCACCTCTCAGGACAGTGGACCTTTACCTGATCGCAACCAATGCGAGGAGA 92017
 QY 1566 GGCACGGCATGTTGGAGTCTCCATCTGCAAGTGGCCACACCTCAGCAACAGCACCCAG 1625
 Db 92018 GGTAGGGCCTGTTAGAGTCTCCATTTGTCAGCTCCACACCTTAGTAACAGCACTAG 92077
 QY 1626 CCGACTGACCCCCCAAGTCCCGCTCTCAGACATCACTGGCTCCAGCAAGCACAGCCG 1685
 Db 92078 CCGATGGCACCCCAAGTCTCGCTTTCGACATCAAGGTTCCAGCAAGCACAGCCG 92137
 QY 1686 GGGAGTGGAGGAGTGGGGGGGAGAGCTTCCAAAAGCCCTCCCGGAAAGGGCTGTGCT 1745
 Db 92138 GGGAGGGAGTGGTGGGGTGGGGGAGCTTCCAAAAGCCCTCCCGGAAAGGGCTGTGCT 92197
 QY 1746 TGTGCTGAAGTACACACACCTCGGCTGGCTGAGTGGTCTGTCAGCAAGTCAAGCACC 1805
 Db 92198 TGTGCTGAAGTATACACACACCTCGGCTGGCTGAGTGGTCTGTCAGCAAGTCAAGCACC 92257
 QY 1806 CCGGTTGAAGTATACACACCTCGGCTGGCTGAGTGGTCTGTCAGCAAGTCAAGCACC 1865
 Db 92258 CAGGTTGAAGTATACACACCTCGGCTGGCTGAGTGGTCTGTCAGCAAGTCAAGCACC 92317
 QY 1866 GATGATCCAGGCTTCCAAAGGCTTCTGTTGTCAGCAAGCCTGGT 1910

Db 92318 GTGACCCAGAGCTACCGTGTGGTTCAGGTGGGACAGGACGACGCT 92362
 RESULT 6
 AC124583/c 303563 bp DNA linear HTG 25-JUN-2002
 LOCUS Mus musculus chromosome UNK clone RP23-123P18, WORKING DRAFT
 DEFINITION SEQUENCE, 14 unordered pieces.
 AC124583
 AC124583.2 GI:21553290
 HTG: HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
 SOURCE house mouse.
 ORGANISM Mus musculus.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 McPherson,J.D. and Waterston,R.H.
 TITLE The sequence of Mus musculus clone
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 303563)
 AUTHORS McPherson,J.D. and Waterston,R.H.
 DIRECT SUBMISSION Direct Submission
 JOURNAL Submitted (14-JUN-2002) Genome Sequencing Center, 4444 Forest Park
 PARKWAY, St. Louis, MO 63108, USA
 REFERENCE 3 (bases 1 to 303563)
 AUTHORS McPherson,J.D. and Waterston,R.H.
 DIRECT SUBMISSION Direct Submission
 JOURNAL Submitted (25-JUN-2002) Genome Sequencing Center, 4444 Forest Park
 PARKWAY, St. Louis, MO 63108, USA
 COMMENT On Jun 25, 2002 this sequence version replaced gi:21426723.

 Center: Washington University Genome Sequencing Center
 Center code: WUGSC
 Web site: http://genome.wustl.edu/gsc/index.shtml
 Contact: submissions@watson.wustl.edu

 Project Information

 Center project name: M_BA0123P18

 Summary Statistics

 Sequencing vector: M13; 0%
 Sequencing vector: plasmid; 100%
 Chemistry: Dye-primer ET; 0% of reads
 Assembly program: Phrap; version 0.990319
 Consensus quality: 297108 bases at least Q40
 Consensus quality: 298758 bases at least Q30
 Consensus quality: 299680 bases at least Q20
 Insert size: 211000; agarose-fp
 Quality coverage: 0.00 in Q20 bases; agarose-fp
 Quality coverage: 10.12 in Q20 bases; sum-of-contigs

 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 14 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.
 * *
 * 1829: contig of 1828 bp in length
 * 1929: gap of unknown length
 * 3717: contig of 1788 bp in length
 * 3817: gap of unknown length
 * 5111: contig of 1294 bp in length
 * 5211: gap of unknown length
 * 9311: contig of 4101 bp in length
 * 9411: gap of unknown length
 * 9412: contig of 6249 bp in length
 * 15661: contig of 5843 bp in length
 * 15761: gap of unknown length
 * 21604: contig of 5843 bp in length
 * 21703: gap of unknown length

Db	275190	TGGCCGCCAAGACTTTGCCAAACATGACAGGGCTGGTGATCTGACCTTCTCCAGGAACAC	275131
Qy	726	CATCAGCCACATCCAGCCCTTTTCCTTTCTGAGCTCGAGAGCTCCGGTCCCTGATCT	785
Db	275130	CATCAGCCATCCAGCCCTTTTCCTTTCTGAGCTCGAGAGCTCCGGTCCCTGATCT	275071
Qy	786	TCACAGCAATCGGTGCACAGCCCTGGGAGGACACCTCCGGGGCTGGTCAACCTGCA	845
Db	275070	TCACAGCAACCGGCTACTAGCTTTGGAGGACACACTCCGGGGTCTGGTCAACCTGCA	275011
Qy	846	GCACCTTATCGTGAACAAACACAGCTGGGGGACGAGATGAGGCTTTTGGAGACTT	905
Db	275010	ACACCTTATCGTGAACAAACACAGCTGGGGGACGAGATGAGGCTTTTGGAGACTT	274951
Qy	906	CTGTGTGACATGGAGGATCTGGACTCTCTCAACAAACCTCAATGGCTGGCGTGGGA	965
Db	274950	CTGTGTGACTTTAGAGGACTGGACTCTCTCAACAAACCTCAATGGCTGGCGTGGGA	274891
Qy	966	CTCCGTGCAGCAGTGTCAACTCCACAGCTGAGCTGGACCACACACCTGCTGGATCA	1025
Db	274890	TTCCTGTACGCGCATGGTCAACTCCACAGCTGAGCTGGACCACAACTGCTGGACCA	274831
Qy	1026	CATCGCCGAGGCACCTTTGAGAGCTGCAGAAACTGGCCCGCTGGAFCTCACCTCAA	1085
Db	274830	CATTGTGTAGGCGACCTTTGAGAGCTGCAGAAACTGGCCCGCTGGAFCTCACCTCAA	274771
Qy	1086	TCGGCTGCAGAACTTCCCGGACCCCATCTTTGGCCGCTCCAGGCTTCGGCTTTCAC	1145
Db	274770	TCGGCTGCAGAACTTCCCGGACCCCATCTTTGGCCGCTCCAGGCTTCGGCTTTCAC	274711
Qy	1146	AGCCACACTTTGCCCCACCTTTGCTTACTTTGGGGTAAACCCACTTCACTGCAA	1205
Db	274710	TGCCACGCTTTGCCCCACCTTTGCTTACTTTGGGGTAAACCCACTTCACTGCAA	274651
Qy	1206	TTGTGAGCTTCTGGGTGGGGAGGCTGCAGCGGGAGTAGACTCGAAACCTTGGCTC	1265
Db	274650	TTGTGAGCTTCTGGGTGGGGAGGCTGCAGCGGGAGTAGACTCGAAACCTTGGCTC	274591
Qy	1266	CCAGGGGGCTCAAGGGTCCGTACTTCTGGCTGCTGGAGGAGGATTTGTGGCGA	1325
Db	274590	CCCGGAAAGTCTCAAGGGTCCGTACTTCTGGCTGCTGGAGGAGGATTTGTGGCGA	274531
Qy	1326	GCCGCTCTCATCACCAGCACACAAAGTCTGGTCTGGTTCGAGCCGAGGGCCAC	1385
Db	274530	GCCGCTCTCATCACCAGCACACAAAGTCTGGTCTGGTTCGAGCCGAGGGCCAC	274471
Qy	1386	ACTCAAGTCAAGCCATGGGGACCCGAGCCCTTATCCACTGGGTAGCCCGGATGA	1445
Db	274470	TCTTAAGTCAAGCCATGGGGACCCGAGCCCTTATCCACTGGGTAGCCCGGATGA	274411
Qy	1446	CCGCTGTAGGAACTCTTCCAGGACTGCTATATGACAAATGGCAATGGCACTTCT	1505
Db	274410	CCGCTGTAGGAACTCTTCCAGGACTGCTATATGACAAATGGCAATGGCACTTCT	274351
Qy	1506	CATCACCACTCAGGACAGTGGTCTTCACTGCAATGCTGCAATGCTGCCGGAGA	1565
Db	274350	AATTAACCACTCAGGACAGTGGTCTTCACTGCAATGCTGCAATGCTGCCGGAGA	274291
Qy	1566	GGCCAGCCCATGGTGGAGGCTTCCATCGTCCAGTGCACACCTCAGCAACAGCAGC	1625
Db	274290	GGCTACCGCCCATGGTGGAGGCTTCCATCGTCCAGTGCACACCTCAGCAACAGCAGC	274231
Qy	1626	CCGCACTGCACCCCGCAAGTCCCGCTCTCAGACATCATGGCTCCAGCAAGACCGG	1685
Db	274230	CCGCACTGCACCCCGCAAGTCCCGCTCTCAGACATCATGGCTCCAGCAAGACCGG	274171
Qy	1686	GGGAGGTTGGAGGCTGGGGGAGGAGCTCCCAAAGCCCGGCAACCGGCTGTCT	1745
Db	274170	GGGAGGTTGGAGGCTGGGGGAGGAGCTCCCAAAGCCCGGCAACCGGCTGTCT	274111
Qy	1746	TGTGTCTGAAGTGCACCCACTCGGCTGTCAAGTGTCTGCTCAGCAAGTCAAGC	1805

*	21704	33356: contig of 11653 bp in length
*	33357	33456: gap of unknown length
*	33457	4470: contig of 11014 bp in length
*	4471	44570: gap of unknown length
*	44571	76854: contig of 32284 bp in length
*	76855	76954: gap of unknown length
*	76955	115914: contig of 38960 bp in length
*	115915	116014: gap of unknown length
*	116015	144525: contig of 28511 bp in length
*	144526	144625: gap of unknown length
*	144626	200283: contig of 55658 bp in length
*	200284	200383: gap of unknown length
*	200384	302135: contig of 101652 bp in length
*	302136	303563: contig of 1428 bp in length.
FEATURES		
source	1..303563	Location/Qualifiers
		/organism="Mus musculus"
		/db_xref="taxon:10090"
		/chromosome="UNK"
		/clone="RP23-123P18"
misc_feature	1..1828	/note="assembly_name:Contig24"
misc_feature	1929..3716	/note="assembly_name:Contig25"
misc_feature	3817..5110	/note="assembly_name:Contig26"
misc_feature	5211..9311	/note="assembly_name:Contig27"
misc_feature	9412..15660	/note="assembly_name:Contig28"
misc_feature	15761..21603	/note="assembly_name:Contig29"
misc_feature	21704..33356	/note="assembly_name:Contig30"
misc_feature	33457..44470	/note="assembly_name:Contig31"
misc_feature	44571..76854	/note="assembly_name:Contig32"
misc_feature	76955..115914	/note="assembly_name:Contig33"
misc_feature	116015..144525	/note="assembly_name:Contig34"
misc_feature	144626..200283	/note="assembly_name:Contig35"
misc_feature	200384..302035	/note="assembly_name:Contig36"
misc_feature	302136..303563	/note="assembly_name:Contig9"
BASE COUNT	81812 a 67733 c 68961 g 83749 t	1308 others
ORIGIN		
Query Match	37.1%	Score 1166.6; DB 2; Length 303563;
Best Local Similarity	86.6%	Pred. No. 1.9e-192;
Matches 1286; Conservative	0;	Mismatches 199; Indels 0; Gaps 0;
Qy	426	GGGTAGATGTCCCTGCCCGGCTGAGTGACCAGACCATGGAGCCCTGCTGGTGG 485
Db	275430	GTGTCTTCTTCCAGTGCAGAGGCTGAGTGACCAGACCATGGAGACTGCTGGTGG 275371
Qy	486	CCTGTAGCTTTGGCATGGCTTTGGCTGATGGCTGATGGCTGATGGCTGATGGCTG 545
Db	275370	GCTGTGGCTTTGGCATGGCTTTGGCTGATGGCTGATGGCTGATGGCTGATGGCTG 275311
Qy	546	CCAGAACTGTGTGACTTGGGACCCCTGTCCTCCAAAGGGGCTGCTTTGTACC 605
Db	275310	CCAGAACTGTGTGACTTGGGACCCCTGTCCTCCAAAGGGGCTGCTTTGTACC 275251
Qy	606	CCCTGATATTGACCCGGGACAGTGGAGCTGCGCTGGGGGCAACTTCATCCACAT 665
Db	275250	CCCTGATATTGACCCGGGACAGTGGAGCTGCGCTGGGGGCAACTTCATCCACAT 275191
Qy	666	CAGCCCGGAGGACTTTGCCAAACATGACGCGGGCTGGTGACCTGTCCAGGAACAC 725

Db 274110 TGTGTCGATGTTACCACACACATCTCCCTGGTCAAGTGGTCTGTCCAGCAAGTCCAGCACC 274051
 Qy 1806 CCGGGTGAAGATGATACCACTGTCAGTCAACTGCTCTGACGATGAGGTACTGATTTACAG 1865
 Db 274050 CAGGGTGAAGATGATACCACTGTCAGTCAACTGCTCTGATGATGAGTCTGATCTACAG 273991
 Qy 1866 GATGATCCAGCCTCAACAGCCTTCGTGGTCAACAACCTGGT 1910
 Db 273990 GTGAGCCAGGCTACCGTGTGGTCAAGTGGGACAGGACCGCT 273946

RESULT 7
 AC125047/c 308013 bp DNA linear HTG 29-JUL-2002
 LOCUS Mus musculus chromosome UNK clone RP23-354G14, WORKING DRAFT
 DEFINITION SEQUENCE, 7 unordered pieces.
 ACCESSION AC125047
 VERSION AC125047.2 GI:22002214
 KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLLTOP.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 308013)
 AUTHORS McPherson, J.D. and Waterston, R.H.
 TITLE The sequence of Mus musculus clone
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 308013)
 AUTHORS McPherson, J.D. and Waterston, R.H.
 TITLE Direct Submission
 JOURNAL Submitted (20-JUN-2002) Genome Sequencing Center, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
 REFERENCE 3 (bases 1 to 308013)
 AUTHORS McPherson, J.D. and Waterston, R.H.
 TITLE Direct Submission
 JOURNAL Submitted (29-JUL-2002) Genome Sequencing Center, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
 COMMENT On Jul 29, 2002 this sequence version replaced gi:21490485.

----- Genome Center -----
 Center: Washington University Genome Sequencing Center
 Center code: WUGSC
 Web site: http://genome.wustl.edu/gsc/index.shtml
 Contact: submissions@watson.wustl.edu
 ----- Project Information -----
 Center project name: M BA0354G14
 ----- Summary Statistics -----
 Sequencing vector: M13; 0%
 Sequencing chemistry: plasmid; 100%
 Chemistry: Dye-terminator; 0% of reads
 Assembly program: Phrap; version 0.990319
 Consensus quality: 303857 bases at least Q40
 Consensus quality: 305426 bases at least Q30
 Consensus quality: 306387 bases at least Q20
 Insert size: 196000; agarose-ftp
 Insert size: 307413; sum-of-contigs
 Quality coverage: 0.00 in Q20 bases; agarose-ftp
 Quality coverage: 9.67 in Q20 bases; sum-of-contigs

 * NOTE: This is a 'working draft' sequence. It currently consists of 7 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.
 * 1 12436: contig of 12436 bp in length
 * 12437 12536: gap of unknown length
 * 12537 40835: contig of 28299 bp in length
 * 40836 40935: gap of unknown length

* 40936 58095: contig of 17160 bp in length
 * 58096 58195: gap of unknown length
 * 58196 142400: contig of 84205 bp in length
 * 142401 142500: gap of unknown length
 * 142501 202912: contig of 60412 bp in length
 * 202913 203012: gap of unknown length
 * 203013 306579: contig of 103567 bp in length
 * 306580 308013: contig of 1334 bp in length.
 * 306580 Location/Qualifiers

source
 1 .308013 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /chromosome="UNK"
 /clone="RP23-354G14"
 1 .12436 /note="assembly_name:Contig10"
 misc_feature 12537..40835 /note="assembly_name:Contig11"
 misc_feature 40936..58095 /note="assembly_name:Contig12"
 misc_feature 58196..142400 /note="assembly_name:Contig13"
 misc_feature 142501..202912 /note="assembly_name:Contig14"
 misc_feature 203013..306579 /note="assembly_name:Contig15"
 misc_feature 306680..308013 /note="assembly_name:Contig18"
 BASE COUNT 90720 a 61748 c 61427 g 93518 t 600 others
 ORIGIN

Query Match 37.1%; Score 1166.6; DB 2; Length 308013;
 Best Local Similarity 86.6%; Pred. No. 1.9e-192;
 Matches 1286; Conservative 0; Mismatches 199; Indels 0; Gaps 0;
 Qy 426 GGGTAGATGTCCTCGCCCGCCCGCCCTGAGTGACACAGACCATGGAGACCCCTGCTGGTGG 485
 Db 26361 GTGTCCTCTTTCCAGGTGCAGAGCTGAGTGACACAGACCATGGAGACTCTGCTGGTGG 26302
 Qy 486 CCGTGTAGCGTTGGCATGGGTTGGCGTGGTGCAGCCCTGCCCAAGTACTGTGCTG 545
 Db 26301 GCTGCTGGCTTTGGCATGGGCTTTGCTGTGCTGCTGCTGCTGCCCAAGTACTGTGCTG 26242
 Qy 546 CCAGAAATCTGTGTGAGTCACTGGGACCCCTGTGCCCTCCAAAGGGGCTCTTTGTACC 605
 Db 26241 CCAGAAATCTGTGTGAGTCACTGGGACCCCTGTGCCCTCCAAAGGGGCTCTTTGTACC 26182
 Qy 606 CCTGATATTTGACCGGGGACAGTGGAGCTGGCCCTGGGGGCAACTTCATCATCCACAT 665
 Db 26181 CCTGATATTTGACCGGGGACAGTGGAGCTGGCCCTGGGGGCAACTTCATCATCCACAT 26122
 Qy 666 CAGCCCGCAGGACTTTGCCACATACCGGGGCTGGGACCTGACCCCTGTCAGGAAACAC 725
 Db 26121 TGGCCGCAAGACTTTGCCCAATGACAGGGGCTGGTGGATCTGACCTGTGCCAGAACAC 26062
 Qy 726 CATCAGCCACATCCAGCCCTTTTCCTTTCTGGACCTCGAGACCTCCGCTCCCTGCATCT 785
 Db 26061 CATCAGCCATATCCAGCCCTTTTCCTTTCTGGACCTCGAGACCTCCGCTCCCTGCAC 26002
 Qy 786 TGACAGCAATCGGCTGCCAAGCCTTGGGAGGACACCCCTCCGGGGCTGGTCAACCTGCA 845
 Db 26001 TGACAGCAATCGGCTGCCAAGCCTTGGGAGGACACACCTCCGGGGCTGGTCAACCTGCA 25942
 Qy 846 GCACCTTATCGTGAACAACAACACCTGGGGCATCCAGATGAGGCTTTTGGAGACTT 905
 Db 25941 ACACCTTATTTGAACAATAACACAGTGGTGGCATCGTGTGATGATGATGATGATGAT 25882
 Qy 906 CCGTGTGACATTTGGAGGATCTGGACTCTCTCCATAACAACACCTCCATGGCCCTCCCTGG 965
 Db 25881 CTTGCTGACTTTAGAAAGACCTGGACCTATCTTTACAACAACCTCCAGGACTTCCCTGG 25822
 Qy 966 CTCCGTCGACGATGGTTCACCTCCACGCTGAGCCCTGGACCAACAACCTGCTGGATCA 1025

ORGANISM

Rattus norvegicus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

REFERENCE
AUTHORS

1 (bases 1 to 197893)
Muzny, D.M., Adams, C., Adio-Oduola, B., Ali-osman, F.R., Allen, C.,
Albrooks, S.L., Amaratunge, H.C., Are, J.R., Ayele, M., Banks, F.,
Barbara, J., Benton, J., Bimage, K., Blankenburg, K., Bonnin, D.,
Barbora, J., Bowie, S., Brieva, M., Brown, E., Brown, M., Bryant, N.P.,
Buhay, C., Burch, P., Burkett, C., Surrall, K.L., Byrd, N.C.,
Carron, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D.,
Chen, G., Chen, R., Chen, Z., Chowdhry, I., Christopoulos, C.,
Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R.,
Davila, M.I., Davis, C., Davy-Carroll, L., Dederich, D.A.,
Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J.,
Earhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M.,
Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P.,
Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R., K.,
Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K.,
Harris, C., Harris, K., Hart, M., Haylak, P., Hawes, A., Hernandez, J.,
Hernandez, O., Hodgson, A., Hogues, M., Holloway, C., Hollins, B.,
Horns, F., Howard, S., Huber, J., Hulyk, S., Hume, J., Jackson, L.E.,
Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S.,
Karleson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C.,
Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L.,
Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W., Louiseged, H.,
Lozado, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J.,
Maneswari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E.,
Massey, E., Mahoney, E., Nwokenwo, S., Ogun, M., Okunolu, G.,
Moser, M., Neale, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N.,
Nguyen, N., Nickerson, E., Nwokenwo, S., Ogun, M., Peery, J., Perez, L.,
Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L.,
Peters, L., Pickens, R., Primus, E., Fu, L.L., Quiles, M., Ren, F.,
Rivers, M., Rojas, A., Rojubokan, I., Rolfe, M., Ruiz, S., Savery, G.,
Scherer, S., Scott, G., Shen, H., Shoohtari, N., Sisson, I.,
Sodergren, E., Sonaake, T., Sparks, A., Stanley, H., Stone, H.,
Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H.,
Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S.,
Usmani, K., Vasquez, L., Vera, V., Villalon, D., Vinson, R., Wang, Q.,
Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S.,
Williams, C., Williamson, A., Wleczyk, R., Wooden, S., Worley, K.,
Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D.,
Weinstock, G. and Gibbs, R.

TITLE

Direct Submission
Unpublished
Worley K.C.

REFERENCE

Submitted (19-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

AUTHORS

Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu

TITLE

Center project name: GXFO
Center clone name: CH230-5008
Summary Statistics

JOURNAL

Sequencing vector: Plasmid;
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 143066 bases at least Q40
Consensus quality: 151342 bases at least Q30
Consensus quality: 157286 bases at least Q20

AUTHORS

* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).

JOURNAL

* NOTE: This is a 'working draft' sequence. It currently
consists of 63 contigs. The true order of the pieces

COMMENT

Table with columns: ID, Label, Sequence, and Description. Contains genomic data for Rattus norvegicus clone CH230-5008, including coordinates, sequence alignments, and accession information.

RESULT 8
AC127880
LOCUS
DEFINITION
AC127880
VERSION
AC127880.1 GI:21908317
KEYWORDS
HTG; HTGS_PHASE1.
SOURCE
Rattus norvegicus.

* is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

- * 1 1082: contig of 1082 bp in length
- * 1083 1182: gap of unknown length
- * 1183 2644: contig of 1462 bp in length
- * 2645 2744: gap of unknown length
- * 2745 3881: contig of 1137 bp in length
- * 3882 3982: gap of unknown length
- * 3983 5344: contig of 1362 bp in length
- * 5345 6742: contig of 1299 bp in length
- * 6743 6843: gap of unknown length
- * 6844 8344: contig of 1501 bp in length
- * 8345 8444: gap of unknown length
- * 8445 10197: contig of 1753 bp in length
- * 10198 11522: contig of 1226 bp in length
- * 11523 13415: contig of 1793 bp in length
- * 13416 13515: gap of unknown length
- * 13516 15170: contig of 1655 bp in length
- * 15171 17014: contig of 1744 bp in length
- * 17015 18688: contig of 1574 bp in length
- * 18689 18788: gap of unknown length
- * 18789 20090: contig of 1302 bp in length
- * 20091 20190: gap of unknown length
- * 20191 22362: contig of 2072 bp in length
- * 22363 24086: contig of 1724 bp in length
- * 24087 24187: gap of unknown length
- * 24188 25869: contig of 1683 bp in length
- * 25870 28059: contig of 2090 bp in length
- * 28060 28159: gap of unknown length
- * 28160 30192: contig of 2033 bp in length
- * 30193 32284: contig of 1932 bp in length
- * 32285 34517: contig of 2193 bp in length
- * 34518 35832: contig of 1215 bp in length
- * 35833 35932: gap of unknown length
- * 35933 38365: contig of 2433 bp in length
- * 38366 39651: contig of 1186 bp in length
- * 39652 42696: contig of 2945 bp in length
- * 42697 44523: contig of 1727 bp in length
- * 44524 47132: contig of 2509 bp in length
- * 47133 49315: contig of 2083 bp in length
- * 49316 49415: gap of unknown length
- * 49416 52479: contig of 2964 bp in length
- * 52480 54484: contig of 2005 bp in length
- * 54485 57588: gap of unknown length
- * 57589 59983: contig of 2295 bp in length
- * 59984 62833: contig of 2750 bp in length
- * 62834 65460: contig of 2527 bp in length
- * 65461 67597: contig of 2037 bp in length
- * 65561

- * 67598 67697: gap of unknown length
- * 67698 71010: contig of 3313 bp in length
- * 71011 73099: contig of 1989 bp in length
- * 73100 73199: gap of unknown length
- * 73200 76559: contig of 3460 bp in length
- * 76560 76660: gap of unknown length
- * 76661 79328: contig of 2569 bp in length
- * 79329 79429: gap of unknown length
- * 79430 82951: contig of 3522 bp in length
- * 82952 83051: gap of unknown length
- * 83052 85337: contig of 2287 bp in length
- * 85338 85338: gap of unknown length
- * 85339 88632: contig of 3195 bp in length
- * 88633 88732: gap of unknown length
- * 88733 92607: contig of 3875 bp in length
- * 92608 92707: gap of unknown length
- * 92708 95823: contig of 3116 bp in length
- * 95824 95923: gap of unknown length
- * 95924 98417: contig of 2494 bp in length
- * 98418 98517: gap of unknown length
- * 98518 102327: contig of 3710 bp in length
- * 102328 102327: gap of unknown length
- * 102328 106214: contig of 3887 bp in length
- * 106215 106314: gap of unknown length
- * 106315 109534: contig of 3220 bp in length
- * 109535 113077: contig of 3443 bp in length
- * 113078 113077: gap of unknown length
- * 113078 116869: contig of 3692 bp in length
- * 116870 116869: gap of unknown length
- * 116871 121793: contig of 4824 bp in length
- * 121794 121893: gap of unknown length
- * 121894 126597: contig of 4704 bp in length
- * 126598 126597: gap of unknown length
- * 126598 131888: contig of 5191 bp in length
- * 131889 131888: gap of unknown length
- * 131889 136453: contig of 4465 bp in length
- * 136454 136553: gap of unknown length
- * 136554 141386: contig of 4833 bp in length
- * 141387 141886: gap of unknown length
- * 141887 146790: contig of 5304 bp in length
- * 146791 146890: gap of unknown length
- * 146891 151544: contig of 4654 bp in length
- * 151545 151644: gap of unknown length
- * 151645 157304: contig of 5660 bp in length
- * 157305 157404: gap of unknown length

Query Match 36.6%; Score 1151.4; DB 2; Length 197893;
 Best Local Similarity 86.6%; Pred. No. 8.5e-190;
 Matches 1269; Conservative 0; Mismatches 196; Indels 0; Gaps 0;

QY 426 GGTAGATGCTGCTCCCGCCAGCGCTGAGTACACAGACCATGGAGACCCTCTTGGTGG 485
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 42887 GTGTCTTCTTCCAGGTGCAGAGCCGAGTACCAGACCATGGAGACTCTGCTTGGCGG 42946

QY 486 CCTGTAGGCTTTGGCATGCGTTTTGCCGTGTCAGCGCTGCCCAAGTACTGTGTCTG 545
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 42947 GCTGCTGGCTTTGGCATGCGTTTTGCCGTGTCAGCGCTGCCCAAGTACTGTGTCTG 43006

QY 546 CCAGAACTGTCTGAGTCACTGGGGACCCCTGTGCCCTCCAAAGGGGCTGCTTTTGTACC 605
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 43007 CCAGAACTGTCTGAGTCACTGGGGACCCCTGTGCCCTCCAAAGGGGCTGCTTTTGTACC 43066

QY 606 CCTGTATTTGACCGCGGACAGTGGAGCTGGCCCTGGCGGGAACCTTCAATCAATCCACAT 665
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 43067 CCCTGACATTGACCGGAGGACAGTGGAGCTGGCTTGGGTGGCAACTTCAATCAATCCACAT 43126

QY 666 CAGCCCGCAGGACTTTGCCAAACATGACGGGCTGGTGGGACCTGACCCCTGTCCAGGAACAC 725
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 43127 TGGCCCGCAGGACTTTGCCAAACATGACGGGCTGGTGGGATCTGACCTTGTCTAGGAACAC 43186

QY 726 CATCAGCCACATCCAGCCCTTTTCTTCTGGACCTCGAGAGCCCTCCGCTCCCTGCATCT 785

44267 CAGGGTAAAGATGATATCAGCTGAGTCAACTGCTCCGACGAGGAGTACTGATTACAG 44326

1866 GATGATCCAGCCCTCCCAACAGGCC 1890
44327 GTGAGCCAGAGGCTGCAAGTGTGGTC 44351

AC127880 197893 bp DNA linear HTG 19-JUL-2002
Rattus norvegicus clone CH230-5008, *** SEQUENCING IN PROGRESS ***,
63 unordered pieces.

AC127880 1 GI:21908317
HTG: HTGS_PHASE1
Rattus norvegicus.

Rattus norvegicus
Rattus norvegicus
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

Rattus.

1 (bases 1 to 197893)
Muzny, D.M., Adams, C., Adio-Oduola, B., Ali-osmani, F.R., Allen, C.,
Alisbrooks, S.J., Amaral, H.C., Are, J.R., Ayele, M., Banks, T.,
Barbata, J., Benton, J., Bimaga, K., Blankenburg, K., Bonnin, D.,
Bouck, J., Bowie, S., Brieva, M., Brown, M., Bryant, N.P.,
Buhay, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C.,
Carroll, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D.,
Chen, G., Chen, R., Chen, Z., Chowdhry, I., Christopoulos, C.,
Cleveland, C.D., Cox, C., Coyne, M.D., Dathorne, S.R., David, R.,
Davila, M.L., Davis, C., Davy-Carroll, L., Dederich, D.A.,
Delaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, H.H.,
Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J.,
Earhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M.,
Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P.,
Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R.,
Gorell, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K.,
Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J.,
Hernandez, O., Hodgson, A., Hogues, M., Holloway, C., Hollins, B.E.,
Homsy, P., Howard, S., Huber, J., Hulyk, S., Hume, J., Jackson, L.E.,
Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S.,
Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C.,
Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L.,
Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W., Loulseghe, H.,
Lozashwari, M., Lu, X., Lucier, R., Lucia, R., Luna, R., Martinez, E.,
Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Mei, G., Metzker, M.,
Massey, E., Mawhney, E., McLeod, M.P., Meador, M., Morgan, M., Morris, S.,
Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S.,
Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N.,
Nguyen, N., Nickerson, E., Nwokenkwo, S., Ogih, M., Okwuonu, G.,
Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L.,
Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y.,
Rives, M., Rojas, A., Rojibokan, I., Rolfe, M., Ruiz, S., Savery, G.,
Scheer, S., Scott, G., Shen, H., Shoostari, N., Sison, I.,
Sodergren, E., Sonaite, T., Sparks, A., Stanley, H., Stone, H.,
Sutton, A., Svatek, A., Tabot, P., Tamerisa, A., Tamerisa, K., Tang, H.,
Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S.,
Umami, K., Vasquez, L., Vera, V., Villalon, D., Vinson, R., Wang, Q.,
Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S.,
Williams, G., Williamson, A., Wleczky, R., Wooden, S., Worley, K.,
Wu, C., Wu, Y., Zhou, J., Zorrilla, S., Nelson, D.,
Weinstock, G. and Gibbs, R.

Direct Submission
Unpublished
2 (bases 1 to 197893)
Worley, K.C.

Direct Submission
Submitted (19-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/

43187 CATCAGCCATCCAGCCCTCTCTCTCTGACCTTGAGACCTTGAGACCTCCGTTCCCTGACCT 43246

786 TGACAGCAATCGCTGCCAAGCCCTTGGGAGGACACCTCCGGCCCTGGTCAACCTGCA 845
43247 TGACAGCAACCGCTACCTAGCTTGGGAGGACACCTCCGGCCCTGGTCAACCTGCA 43306

846 GCACCTATCGTCAACAACAACAGCTGGGGCCGACATCGAGATGAGGCTTTTGAAGACTT 905
43307 GCACCTATTTGGAACAATATCAGCTGGGTGGCAATGCTGATGATGCCCTTTGAAGACTT 43366

906 CTTGCTGACATTTGGAGGATCTGGACCTCTTCAACAACCTCCATGCTCCGCTGGGAGG 965
43367 CTTGCTGACATTTGGAAGATTTGGACCTATCTTCAACAACCTCCAGGCTTCCCTGGGA 43426

966 CTCGCTGACGATGTTCAACCTCCACAGCTGGACCTGGACCAACCTCTGCTGATCA 1025
43427 TTTGTAAGGGCCATGTTCAACCTCCATCAGCTGAGTCTCGACCAACCTTCTGGACCA 43486

1026 CATCCCGAGGACCTTTGACAGCTGCAAACTGGCCCGCTGATCTCACCTCCAA 1085
43487 CATCGCTGAGGATACCTTGGCAGACCTGCAAACTGGCCCGCTGACCTCACGTCCAA 43546

1086 TGGCTGACAGGCTGCCCCGATCCATCTTTGGCCGCTCCAGGCTTGGCTTTCAC 1145
43547 TGGCTGACAGGCTTCCCCAGACCCCACTTTGGCCGCTCCAGGCTTGGCTTTCAC 43606

1146 AGCCACACCTTTGCCCAACCTTGTCTTTAGTTTGGGGTAAACCTTCACTGCAAA 1205
43607 TGGCACACCTTTGCCCAACCTTGTCTTTAGTTTGGGGTAAACCTTCACTGCAAA 43666

1206 TTGTGAGCTTCTCGCTGGGAGGCTCGAGGGGAGATGACCTGGAAACCTTGGCTC 1265
43667 TTGTGAGCTTCTCGCTGGGAGGCTCGAGGGGAGATGACCTGGAAACCTTGGCTC 43726

1266 CCCAGGGGCTCAAGGTCGCTTCTTGGCATGTCGTCGAGGAGGTTTGTGTCGA 1325
43727 CCCTGGAGTCTCAAGGTCGCTTCTTGGCATGTCGTCGAGGAGGTTTGTGTCGA 43786

1326 GCGCTCTCATCCACGACACACACAAAGTTGCTGTTCTGGAGGCGCCAGGGCCAC 1395
43787 GCGCTCTCATCCACGACACACAAAGTTGCTGTTCTGGAGGCGCCAGGGCCAC 43846

1386 ACTCAAGTGAAGCCATTTGGGACCCCGCTTATCCACTGGTATGCCCCGATGA 1445
43847 TCTTAAGTGAAGCCATTTGGGACCCCGCTTATCCACTGGTATGCCCCGATGA 43906

1446 CCGCTGTAGGAACTCTCAAGGACCGCTGTCTATGACAAATGGCACCCCTGGACATCTT 1505
43907 CCGCTGTAGGAACTCTCAAGGACCGCTGTCTATGACAAATGGCACCCCTGGACATCTT 43966

1506 CATCACACATCTCAGGACAGTGTGCTTCACTGCAATGCTGCCAATGCTGCCGGAGA 1565
43967 AATTACACCTCTCAGGACAGGCGGCTTCACTGCAATGCTGCCAATGCTGCCGGAGA 44026

1566 GCGCACGCTTGTGGAGGCTCATCTGCTCCAGCTCCACACTCAGCAACAGCACAGCAG 1625
44027 GCGCACGCTTGTGGAGGCTCATCTGCTCCAGCTCCACACTCAGCAACAGCACAGCAG 44086

1626 CCGCACTGACCCCGCTTCTAGACATCACTGGCTCCAGCAAGACAGGCGG 1685
44087 CCGGATGGCGCCGCTTCTAGACATCACTGGCTTCCAGCAAGACAGGCGG 44146

1686 GGGAGGTGAGGAGTGGGGGGAGAGCTCCCAAAAGCCCGGAAACCGGCTGTGCT 1745
44147 GGGAGGTGAGGAGTGGGGGGAGAGCTCCCAAAAGCCCGGAAACCGGCTGTGCT 44206

1746 TGTGCTGAAGTGAACCAACCTTGGCCCTGTCAAGTGTCTGTCACCAAGTCAAGCACC 1805
44207 TGTGCTGAAGTGAACCAACCTTGGCCCTGTCAAGTGTCTGTCACCAAGTCAAGCACC 44266

1806 CCGGTCAGATGTACCAGTGCAGTACAACTGCTGACGATGAGGACTGATTTACAG 1865

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GXPO
Center clone name: CH230-5008
----- Summary Statistics

Sequencing vector: Plasmid
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 143066 bases at least Q40
Consensus quality: 151342 bases at least Q30
Consensus quality: 157286 bases at least Q20

* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
consists of 63 contigs. The true order of the pieces
is not known and their order in this sequence record is
arbitrary. Gaps between the contigs are represented as
runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
as soon as it is available and the accession number will
be preserved.

Table with 2 columns: Accession number and Description. Accession numbers range from 1082 to 157305. Descriptions include contig lengths and gap sizes.

Query Match 26.3%; Score 827.4; DB 2; Length 197893;
Best Local Similarity 80.9%; Pred. No. 1,1e-133;
Matches 1062; Conservative 0; Mismatches 216; Indels 35; Gaps 7;

QY 1859 TTTACAGATGATCCAGCCTCCAAACAGGCCCTTGGTGTGACACACCTGTCAGGGA 1918
Db 194803 TCTGAGGATGATCCAGCCTCCAAACAGGCCCTTGGTGTGACACACCTGTCAGGGA 194744
QY 1919 CTGGTACGACTGTGTGTGTGGCCATGTGGGATGACACACCCAGCACCTCAGGGCCA 1978

```

Db 193669 ACCCCCTTCCCTATGTACCTTCCAGGCCCTTCCCACCCAGCCCGGGGTGTCTCAG 193610
Qy 3032 GGAATGTGACTCGCTCAAATCGGACTGAGCCCTGAGTGTGGAAAGGCGAGACTCC 3091
Db 193609 GGAATG-GGACTCGCTCATGTGTCTGACTGAGCCCTGAGTGTGGAAAGGCTAACAACT 193551
Qy 3092 GCCTTCTTAATCAAAAATGTAGCCTTCAAAGCAAGCGGCTTGGATTGCTTATG 3144
Db 193550 GCCTTCTTAATCAAAAATGTAG-CGACAAGCAAGCGGCTTGGATTGCTTATG 193499

RESULT 10
AX179258 1887 bp DNA linear PAT 03-JUL-2001
LOCUS Sequence 3 from Patent WO0142471.
DEFINITION AX179258
ACCESSION AX179258
VERSION AX179258.1 GI:14598947
KEYWORDS synthetic construct.
SOURCE synthetic construct.
ORGANISM artificial sequences.
REFERENCE 1 (bases 1 to 1887)
AUTHORS Prayaga, S.K. and Shimkets, R.A.
TITLE Polypeptides and polynucleotides encoding same
JOURNAL Patent: WO 0142471-A 3 14-JUN-2001;
Curagen Corporation (US)
FEATURES
Location/Qualifiers
Source
1..1887
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="Curagen clone AF038458_A"

BASE COUNT 259 a 736 c 577 g 315 t
ORIGIN
Query: Match 20.2%; Score 634.6; DB 6; Length 1887;
Best Local Similarity 61.3%; Pred. No. 4.1e-100;
Matches 1079; Conservative 0; Mismatches 664; Indels 16; Gaps 3;
Qy 507 GTTTCCTGCTGAGCGCTCCCAAGGGCTGTCTTTGTACCCCTGATATTGACCGCGGAC 626
Db 63 GTCAGCCAGACCCAGCCCATGTCCCGCGCTGCGCTGCCAGACACAGTCTGTCGCCCT 122
Qy 567 GGGGACCCTGTGCCCTCAAAGGGCTGTCTTTGTACCCCTGATATTGACCGCGGAC 626
Db 123 AAGCGTGTGTGTCAGGGGAGGCTCTCTTGTGTCACCCCTCGCTGGACCGCGGAC 182
Qy 627 AGTGGAGCTGCGCTTGCGCGGCAATTTCATCATCCACATCAGCCCGCAGGACTTGGCAA 686
Db 183 AGCCGAGCTGCGCTGGCAGCAAACTTTCATCCCTCCGTGCGCCGCGGACCTGGCCAA 242
Qy 687 CATGACGGGCTGTGGACTGACCTGTCCAGGAACACCATCAGCCACATCCAGCCCTT 746
Db 243 CATGACAGGCTGTGCTGATCTGAGCTGTTCGGGAAACCATCCGCGCGCTGGTCCCG 302
Qy 747 TCCCTTTCGACCTCGAGGCTCCGCTCCCTGCACTTTGACAGACTTCCCTGACATTCGCA 806
Db 303 CGCTTTCGCGACTCGGGCCCTGCGGCTCCCTGCACTTTGACAGACTTCCCTGACATTC 362
Qy 807 CTTTGGGGAGGACACCCCTCCGGGCTGTGCTCAACCTTCAGCTGAGCACTTATCGTGAACA 866
Db 363 ACTGGGCGAGGGCCAGCTGCGGGCTGTGCAACTTTCGCGCCACCTCATCTTCAGCAACA 422
Qy 867 CCGAGCTGGGGGATCGAGATGAGGCTTTTGGAGACTTCCCTGCTGACATTCGAGATCT 926
Db 423 CCGAGCTGGAGGCTGGCGGCTGGGCTCCCTGGATGATTTGTCCGAGACACTGGAGGACCT 482
Qy 927 GGAGCTCTCCTCAACAACCTCCATGCGCTGGGACTCGTGGGAGCTCGTGGGAGCTGCTCAA 986
Db 483 CGAGCTCTCCTCAACAACCTCCATGCGGAGCTGGCTGGAGGCTGGGAGGCTGGGCTGGCAA 542
Qy 987 CCTTCAAGCTGAGCTGGACACCAACCTGTGGATTCACATCCCGGAGGCGGACCTTTTC 1046
Db 543 CGTCAAACGTTGGGCTTCGACCAACACTTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 602

```

```

Db 194743 CGGGCTACGACTTGTGGTACTTGGCTATGTTGGGAGACACAGCCCAACCGTCTACGGCCA 194684
Qy 1979 CCAACATCGTGGCTCGCCCAAGTTCACCAAGGCTGACTTACCAGGCTGGCAGTCCCAAGTCCA 2038
Db 194683 CCAACATCGTGGCTCGCCCAAGTTCACCAAGGCTGACTTACCAGGCTGGCAGTCCCAAGTCCA 194624
Qy 2039 TGCACAGCCAGATTCTGGGGGGCACCATGATCCTGGTCAATCGGGGGCAATCAATCGTGGCCA 2098
Db 194623 TGCACAGCCAGATTCTGGGGGGCACCATGATCCTGGTCAATCGGGGGCAATCAATCGTGGCCA 194564
Qy 2099 CGCTGTGCTTTCATGTCATCTCATGTTGGTGTACAGGCTTCCAAACAGAGGCGCC 2158
Db 194563 CACTGTGCTTTCATGTCATCTCATGTTGGTGTACAGGCTTCCAAACAGAGGCGCC 194504
Qy 2159 CCAGCAAGATGGAGGGCC---GTGAGCAATGTGTACTCGAGACCAACAGCGGCCAGC 2215
Db 194503 CAGGCAAAATGGAGGGCCACTGTGAGCAATGTGTACTCGAGACCAACAGCGGCCAGC 194444
Qy 2216 CACCGCTCCCAAGCAGCCACCAAGCGGGGCCCGCGGAGGCGCCGCGAAGGTTGGTGG 2275
Db 194443 CTCACCTCTGGGTGGGATGCTTGTGGCCAGCTTCTCAGGCACCAACCGAGGCTGGTGG 194384
Qy 2276 TGGCAACAGACTCTTGGACTTACCAGCAGCTGGCCGCGCAGTGTACTCTCTTCTCT 2335
Db 194383 TGGCAATGATGATGGACTTACGTACCAAGCTTGGCAAGAGCTGTGACTTCTCTCT 194324
Qy 2336 CCAGCTCCCTGGGAGTGGGAGGCTCGGGGCTGGGAGCGGGCCCTGGAGGATCCCAAC 2395
Db 194323 CCAGTTCCTGGGAGTGGGAGGCTCAGGGCTGAGCCGGGCTGAGCCGGGCTCCCAAC 194264
Qy 2396 CCTCCGGCCCGCCCAAGCCAGCCTTGAACCGCTGATGGGGCTTGGCCCTCCCTGG 2455
Db 194263 CCCCAGCTCTCTGCCAAGCCAGCCTTGAACCGCTTGTGGGAGCTTTTGGCCCTCCCTGG 194204
Qy 2456 ACCTCAAGCTCAGAAAGAGAGGCTGTGGACTCCAGGACTCCAGCGGGGAGAGGG 2515
Db 194203 ACCTCAAGCTCAGAAAGAGAGACTTTAGACTCTAGGACTCCAGAGGCGAGAGGG 194144
Qy 2516 CTGGGAGCTCGGGCCGGCCCAACCTCGGACCGAGAGGCTGCTGGGGCCCTCCCTGG 2575
Db 194143 CAGGGACATCAGCCAGGGCCCAACCTCAGACCGAGAGGCACTGCTGGGGCCCGCAACCA 194084
Qy 2576 CCGGGCCAGAGCTGCTCCCTTGCCTGGGAGGCAAGCCAAACGAGCCACTCTCT 2635
Db 194083 CCGGTGCCAGGAGCTTCTCCCTTGGCCCTGGAGGCAAGCCAAACGAGGCACTCTT 194024
Qy 2636 TGCATGCGGGGACTTGTCTGTGCGGGCGGGAGGGGCTGTGCGGGGGCTACAGTC 2695
Db 194023 TTGCATGGGGACTTTGCAGCTGCAGCTGCAG-----CTGCTCTCTGGTGGCTACAGTC 193970
Qy 2696 CTCCTCGGAGGTCTCGAATCTGGAAGAGGCTGCTCTGTCGAAGGCTGCTCT 2755
Db 193969 CCCCTAGGGGCTCTCAAACTCTGGAAGAGGAGGCTGCTGCTCAANTGGATGCTCT 193910
Qy 2756 TGCCTTTGAGGAGTACCTGTGGGGCGGGGGACATTTGGAGCTCCGAAATGGG 2815
Db 193909 TGCCTTTGAGGAGTACCTGTGGGGCGGGGGAGCTTGGAGGCAATTTGGAGCTCAGAGTGGG 193850
Qy 2816 TGAATGAGAGCCGCTTAGTGGGGTGGGCAATGCTCCCTTTCTGTCGGCAGGGTGGG 2875
Db 193849 TAATGAGAGTACTGTGATAGCAGGAGGAGAGTACCCTCCCTCCCGCCCTTCTCAAAGT 193790
Qy 2876 AGNAGGGGAAGATC-----TCACTGGCAAGTGTGTTGGAGTTTCCATGTTGATGT 2928
Db 193789 GGGAGGAGAGGAGTCAAGTCAAGTCAAGTATTTGTGAGTTCATGTTGATGT 193730
Qy 2929 TTACATCCAGGAGACTTTCCTCTCCCTGCAATGAGGCTCTGCTGTC-----CCCCCTAC 2982
Db 193729 TTACATCCAGGAGACTTTCCTCTCCCTGCAATGAGGCTCTGCTGTC-----CCCCAGCTAC 193670
Qy 2983 CCGCAACACCCACATCCTCCCA-----CCACCAGGCGGGGTTGCTCAG 3031

```


1355	CTCACCACTTGGCTGTCGCCGACAGTTCGGCCGCTGCCCTGCGCTGCCGGCAGTGGG	1414
1407	GGACCCAGCCCTTATCCACTGGGTAGCCCTCCCGATGACCCCGTGGTAGGAACTCCTC	1466
1415	GGACCCAGACCCCGTGGTGGTGTACCCAGGCGCGGTCTAGGCAACTCAAG	1474
1467	AAGGACCGCTGTCTATGACAAATGGACCCCTGGACATTTTCATCCACATCTCAGGACAG	1526
1475	CCGTGCGCGGCTTCCCAATGGACCGTGGAGTCTGGTCCCGAGCCGGGTGATGG	1534
1527	TGFTGCTTCACTGTGATCTCCCAATGGACCGTGGAGTCTGGTCCCGAGCCGGGTGAGGT	1586
1535	TGGCATTTCCACTGATTCGGCCCAATGACGTGGGGAGGCGCACAGTCTGCTGTGAGCT	1594
1587	CTCCATCGTCCAGTGCACACCTCAGCAACACAGCCAGCCGCA--CTGCACCCCAAG	1644
1595	GACTGTGGGTCCCCACCTCTCTAGCTAGCAACAGCACAGCTGTACCCCGCGG	1654
1645	TCCCGCTCTCAGACATCACTGGTCCAGCAAGACCCAGCCGGGAGGTGAGGCACTGGG	1704
1655	GGACGGGATCTTGATCTCTCACCCCACTCCGCTGCTCTCTTCCCAAGGTGGC	1714
1705	GCCGGAGGCTCCCAAGGCCCGGAAAGCGGCTGTGTGTCTGAAAGTACCAACC	1764
1715	CGACACTGGGCGCCCTA-----CCGACCGTGGCGTCCAGTGACTGAGACGGGGCC	1766
1765	ACCTCGGCGCTGTCAGTGTGTGTGTCAGCAAGTACAGCACCCCGGGTGAAGATGTACAG	1824
1767	ACAGTCTTGTTCAGTGGCGGATCAGCGGCTATCCCGGCATCCGATGTACAG	1826
1825	CTGAGTACAACACTGTGTGACGATGAGTACTGATTTACAGGATGATCCCAAGCTCCAAC	1884
1827	ATCCAGTACAACAGCTGGTGTGACATCTCTGTCTACAGGATGATCCCGCGGAGGC	1886
1885	AAGGCTTCTGGTCAACAACTGGTGTGAGGACTGGCTACGACTTGTGTGTGTGGCC	1944
1887	CGTCTGTTCTGTGACGACTGGGCTGAGCGGACCTACGATCTGTGGTGTCTGCC	1946
1945	ATGTGGGATGACACAGCCAGACACTACCGCCACCAACATCTGGGCTGGCCAGTTC	2004
1947	GTGTATGAGACAGCGCCACCGGGCTACCGCCACCGGGCTGTGGGTGGCCCGCTTC	2006
2005	TTCCAAAGCTGACTACCGGAGTCCAGTCCATGTACACAGCCAGATTTCTGGCGGCACC	2064
2007	TCCACCGAACCCTGGCCATGCGGGGCGCCGACCGCTCCCTTCTTGGGCGGACG	2066
2065	ATGATCTTGGTCACTGGGGGATCATCTGTGGCCAGCTGTGGTCTTATCTGTCTTC	2124
2067	ATGATCTTGGTCACTGGGGGATCATCTGTGGCCAGCTGTGGTCTTATCTGTCTTC	2126
2125	ATGGTGGCTTACAGTCTGCAACCCAGGCGCCCGAG-----CAAGATGCGAGCGGCC	2178
2127	CTAATCGGTACAAGGTGACCGGGCCAGCCCGCCCGGCAAGGCCAAGATTCGGCGCT	2186
2179	GTGAGCAATGTGACTCGCAGACCAACCGCGCCCGACCCCGCTTCAAGAGCGGACCA	2238
2187	GTTAGCAGGTTTCTTCCAGACCAACCGCGCCCTGGGCGCCACCGCCACCGCCGCGG	2246
2239	GCCGGGCGCCCGCGCAGG	2257
2247	CCCGCCCGGAGCCCGCGG	2265

507	GTTTCCGCTGCGAGCCCTGCCCAAGGGCTGCTTTTGTACCCCTGATTTGACCGGCTACT	566
515	GTACGCCACACCAGCCCAATGTCGCCGCTGCCCTGCCAGACACAGTCCGTCGCCCT	574
567	GGGACCCCTGTGCCCTCCAAAGGGCTGCTTTTGTACCCCTGATTTGACCGGCTACT	626
575	AAGGCTGCTGTCCAGGGCAGCCCTCTGTTGTCGCCCACTCCGCTGCGACCGCGGGC	634
627	AGTGGAGCTGCGTGGCGGCACTTCATCCATCCACATCAGCGCCAGGACTTTGCCAA	686
635	AGCCGAGCTGCGTGGCAGACAACTTCATGCGCTCCGTGCGCCCGCGGACTGGCCAA	694
687	CATGACGGGCTGTGGACCTGACCTGTCCAGGAACACCATCAGCCCAATCCAGCCCTT	746
695	CATGACGAGCTGTGATCTGAGCTGTGCGGMAACCATCCGCCACCTGGCTGCCGG	754
747	TTCCCTTTTGACCTGAGAGCTTCCCTGCTGATCTTTCAGAGCAATCGGCTGCCAAG	806
755	CGCCTTCCCGACCTGGCGGGCCCTGCTGCTGACCTGGATGGCAACCGGCTGACCTC	814
807	CTTTGGGAGACACCTTCGGGCGCTGTTCAAACCTGACGACCTTATCTGTGAACAA	866
815	ACTGGGAGGGCCAGCTGCGCGCTGTGTAACCTGGGCACTCATCTCAGCAACA	874
867	CCAGTGGCGGCTGCGAGATGAGGCTTTGAGGACTTCTGTGTGACATGAGGATCT	926
875	CCAGTGGCAGGCTGGCGGGCGCGCTGGATGATTTGTGCCAGACACTGGAGGACT	934
927	GGACTTCTTACAACACTTCCATGGCTGCGTGGGACTTCCGTGGGAGCATGGTCAA	986
935	CGACCTCTCTACAACACTTCCATGGCTGCGGCGCTGGAGGCTTGGCGGCTGGCAA	994
987	CTCCACCAAGCTGAGCTGGACCAACACTTCTGATCACAATCGCCGAGGACCTTTTC	1046
995	CGTCAACACTTGGGCTGACCAACACTTCTGCTGGCTTCTGTGCCCGCGGCTTTTC	1054
1047	AGACTGAGAACTGGCGGCTGGATCTCACCTCCAAATCGGCTGCGAGAGCTGCCCC	1106
1055	CGCCCTGCAACAGCTGGCCGCTGGACATGACCTTCAACCGCTTGCACCAATCCAC	1114
1107	TGATCCCACTTTGGCGCTCCAGGCTTGGCTTGGACTTGCAGCCACACCTTTGCCACC	1166
1115	CGACCCACTTCTTCCCGCTGCCCTGCTGCCAGGCGCCCGGGCTGCCCGCTCTGC	1174
1167	CTTTGCTTTTAGTTTGGGGTAAACCACTTCACTGCAATTTGTGAGTCTTCTGGCTGG	1226
1175	CTGTGTGCTGGCTTTGGCGGAAACCCCTGCACTGCAACTGCGAGCTGGTGTGGCTGG	1234
1227	GAGGCTGAGCGGAGTGAACCTGGAACCTGCTGCCCGCGGGGCTCAAGGGTGG	1286
1235	TGCGCTGGCGGGAGGAGACTTCCAGGCTTCCGCTTCCCACTGCTTGGGGGGCG	1294
1287	CTACTTCTGGCATGCTGGAGAGGATTTGTGTCGAGCGGCTTCTCATCCACCAGA	1346
1295	CTACTTCTGGCGGTGGCGAGGAGATTTGTCTGAGGCGCCCGCTGTGACTCACCG	1354
1347	CACACAAGTCTGCTGTGGAGGCGCCAGGGCCACACTCAAGTGCAGGACCAATGG	1406

RESULT 12
 AX266997 LOCUS
 DEFINITION Sequence 3 from Patent WO0172827.
 ACCESSION AX266997
 VERSION AX266997.1
 KEYWORDS human.
 SOURCE Homo sapiens
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Db Qy
 1887 bp DNA linear PAT 26-OCT-2001

Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE 1

Glucksmann, M.A. 33395, a novel human leucine-rich repeat family member and use thereof

JOURNAL Patent: WO 0172827-A 3 04-OCT-2001;

Millennium Pharmaceuticals, Inc. (US)

Location/Qualifiers

1. 1887

/organism="Homo sapiens"

/db_xref="taxon:9606"

257 a 735 c 579 g 316 t

BASE COUNT

ORIGIN

Query Match 20.1%; Score 633; DB 6; Length 1887;

Best Local Similarity 61.3%; Pred. No. 7.7e-100;

Matches 1078; Conservative 0; Mismatches 665; Indels 16; Gaps 3;

QY	507	GTITGCCGTGGCGCCCTGCCCAAGGCTGCTTTGTACCCTGATATTGACCGGGGAC	566
DB	63	GTACGCCACACCAGCCCATGCCCCCGCTGCCAGACACAGTCGTCGCCCT	122
QY	567	GGGACCCCTGTCCTTCCAAAGGCTGCTTTGTACCCTGATATTGACCGGGGAC	626
DB	123	AAGCGTGTGTGCCAGGGGAGCCCTCTCTGTTCGTCACCCTGCTGGACCCGGGGC	182
QY	627	AGTGAGCTCGCCCTGGCGGCAACTTTCATATCCATCCACATCAGCCGCGAGCTTTGCCAA	686
DB	183	AGCCGAGCTCGCGCTGGCAGACAACTTCATCGCCCTCGTGGCCGCGGACCTGGCCAA	242
QY	687	CATGACGGGCTGTGGACCTGACTACCCTGTCCAGGAAACACATCAGCCCAATCAGGCC	746
DB	243	CATGACAGGCTGTGCATGTAGCCTGTCCGGAACACATCCGCGACGTGGCTGCCGG	302
QY	747	TTCCTTTGACCTCGAGNCCTCGCTCCCTGATTTGACAGCAATCGGCTGCCAAG	806
DB	303	CGCCTTTCGGCCACTCGGGGCTCGCTGCTGTCACCTGGATGGCAACCGGCTGACCTC	362
QY	807	CTTTGGGAGGACACCCCTCCGGCCCTGGTCAACCTGCAGACCTTATCGTGAACAACAA	866
DB	363	ACTGGGAGGGCCAGCTGGCGGCTGGTCAACTTTGGCCACCTCATCTCANGAACAA	422
QY	867	CCAGCTGGCGGCTCGCAGATGAGGCTTTTGGAGACTTCCTGCTGACATTTGAGGATCT	926
DB	423	CCAGCTGGCAGGCTGGCGGCTGGATGATTTGCGAGACACTGGAGGACCT	482
QY	927	GGACCTCTCAACAACTCCATGGCTGGCTGGGACTCCGTCGAGCGATGCTCAA	986
DB	483	CGACCTCTCTCAACAACTCGAGCAGTGGCCCTGGGAGGCTGGCGGCTGGGGCAA	542
QY	987	CTTCCACAGCTGAGCTGGACCAACACTGCTGGATCACATCGCCGAGGGCACCTTTGG	1046
DB	543	CGTCAACAGTGGGCTCGACCAACACTGCTGGCTTTGTCGCCGCGGCGCTTTTC	602
QY	1047	AGACCTGCGAATACTGGCGGCTGGATCTACCTCCAATFGGCTCGAAGCTGCCCC	1106
DB	603	CGCCCTGCAACAGCTGGCGGCTGGACATGACTCCAACCGGCTGACCAACATCCACC	662
QY	1107	TGATCCCATTTTGGCCGCTCCAGGCTTGGGTTTTCAGACCCACACCTTTGGCCACC	1166
DB	663	CGACCCACTCTTCTCCGCTGCTCCGCTGGCTGGAGGCTGGCGGCTGGCCCTCTGC	722
QY	1167	CTTGTCTTTAGTTTGGGGTAACCCACTTCACTGCAATTGAGCTTCTCTGGGTGGC	1226
DB	723	CTTGTGTGGCTTTGGGAGGAAACCCCTGCACTGCAACTGGAGCTGGTGTGGCTGG	782
QY	1227	GAGGCTCGAGCGGGACGATGACCTGGAAACCTGTGGCTCCCAAGGGGCTCAAGGGTGG	1286
DB	783	TGCGCTGGCGGGAGGACGACTCGAGGCTCGCGCTCCCAACCTGCTGGCGGGCGG	842
QY	1287	CTACTTCTGGCATGCTGGTAGGAGGATTTGTCGGAGCCGCTCTCATCCACCAGCA	1346
DB	843	CTACTTTGGCGGGTAGGAGGATTTGTCTCGAGCCCGCTGGTGTACTCACCG	902

QY	1347	CACACAAAGTTGCTGTTTGGAGGCGGCGGCACACTCAAGTCAAAGCAAGCATTTGG	1406
DB	903	CTACACACTTGGTGTGCCCGCAGGTGGCGGGTCCCTCCCTCCCGGCAAGTGGG	962
QY	1407	GGACCCAGCCCTTATCCACTGGTAGCCCGGATACCCGCTGGTAGGAACTCCTC	1466
DB	963	GGACCCAGAGCCCGTGGTGGTGTACCCCGGCGGCTGCTAGGCAATCAAG	1022
QY	1467	AAGGACCGCTGTATGACAAATGGCACCCTGGACATCTTTCACACACATCTCAGGACAG	1526
DB	1023	CCGTGCCCGGCTTCCCAATGGACGCTGGAGTGTGGTCAACCGAGCCGGTGTGG	1082
QY	1527	TGTTGCTTACCTGCAATGTTGCATGTCGCGGAGGAGCCAGCCATGGTAGGAT	1586
DB	1083	TGGCATTTACCTGCAATGGCGCAATGCCAGTGGGAGGCCACAGTGTGTGGAGCT	1142
QY	1587	CTCCATGCTCCAGCTGCCACACCTCAGCAACAGCACCAGCCGCA--CTGACCC	1644
DB	1143	GACTGTGGTCCCCCACCCTCTCAGCTAGCCAAACAGCACCACTGTGACCCCGGG	1202
QY	1645	TCCCGCTCTCAGACATCACTGGTCCAGCAAGACAGCCGGGAGGTGGAGGATGGG	1704
DB	1203	GGACGGGATCTGTATGCTTCAACCCACCTCCGCTGCTGCTTCTGCCAAGTGGC	1262
QY	1705	GGCGAGGAGCCCTCCAAAGAGCCCCCGAAAGCGGCTGTCTTGTGTGAAGTACCACC	1764
DB	1263	CGACACTGGGCCCTTA-----CCGACCGTGGCGTCCAGGTGACTGAGCACGGGGC	1314
QY	1765	ACCTCGGCCCTGGTCAAGTGTGTCAGAAAGTCAAGCACCCTGGTGTGNAAGTGTACCAG	1824
DB	1315	ACAGTGTCTTGTCCAGTGGCCGATCAGCGGCTATCCCGGGATCCGCAATGTACCAG	1374
QY	1825	CTGAGTACAACCTGCTGACGATGAGTACTGATTTACAGATATCCAGCCTCCAAAC	1884
DB	1375	ATCCAGTACAACAGTCTGGTGTGATGACATCTCTGCTACAGAGTATCCCGGGAGAGC	1434
QY	1885	AAGGCCCTTGTGGTCAACAACTGGTGTGAGGAGTGGCTACGACTTGTGTGTGGCC	1944
DB	1435	CGCTGTTCTGCTGACGACCTGGCGTCAGCCCGGACTACGATCTGTGCTGTCTGGC	1494
QY	1945	ATGTTGGATGACAGCACAGCACACTCAGCCGACCAACATCGTGGGCTGCCCGC	2004
DB	1495	GTGTATAGGGCAGTGGCCAGGGCTCAGCGGCTGTGGGCTGTGGGCTGCCCGCTTC	1554
QY	2005	TTCCACAAAGGCTGACTACCCGAGTCCAGTCCATGCACAGCCAGATCTGGCGGCACC	2064
DB	1555	TCCACCGAACCTCCGCTGGGCAATGGGGGGCCGACGCTCCCTTCTGGGGGGCAGC	1614
QY	2065	ATGATCTCTGATCATCGGGGCAATCATGTTGGCCACGCTGTGGTCTTTCATCTATCTC	2124
DB	1615	ATGATCATCGCGTGGCGGCGTCACTGTAGCTCGGTACTGCTTTCATCTTCTGTGTG	1674
QY	2125	ATGTTGGCTACAAGGCTGTCAACACAGAGGCGCCCGAG-----CAAGTGCAGGGCC	2178
DB	1675	CTAATGGCTACAAGGTGCAAGGGCGGCGCCCGCCCGGCAAGGCAAGATTC	1734
QY	2179	GTGACAATGTCTACTCCAGAACACCGCGCCACCGCCACCGCTCCAAAGCAGCCACA	2238
DB	1735	GTTAGAGGTTTGTCTCCAGAACCAAGGCGGCTTGGGCCCCCAGCCACCGCCCGCG	1794
QY	2239	GCCGGGGCCCCCGCCAGG	2257
DB	1795	CCCGCCCGGAGCCCGCGG	1813

RESULT 13

AX266995

LOCUS

DEFINITION

AX266995

ACCESSION

AX266995.1

KEYWORDS

AX266995
Sequence 1 from Patent WO0172827.
AX266995
AX266995.1 GI:16515794

DNA linear PAT 26-OCT-2001
2558 bp

Tue Jun 10 08:57:36 2003

SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1
 AUTHORS Glucksman, M.A.
 TITLE 33395, a novel human leucine-rich repeat family member and use thereof
 JOURNAL Patent: WO 0172827-A 1 04-OCT-2001;
 Millennium Pharmaceuticals, Inc. (US)
 FEATURES
 source Location/Qualifiers
 1 2558
 /organism="Homo sapiens"
 /db_xrefs="taxon:9606"
 /notes="unnamed protein product"
 /codon_start=1
 /protein_id="CAD10239.1"
 /db_xref="GI:16515795"
 /translation="MAILPLLLCLLPLASPPOSATPSPRRRCRCQTSPLSLVLFADRLALRHLDGNSRLTSLGQRLGLVNLHLLSNNQLAALAGALDDCAEQLLEDLDSYNNLEQPLWEALRGNVNTLGDHNLHLASVPAGASFLHKLARLDMDLSEACASPP
 IIPDPLFRPLRLPRGSPASALVAFAGNPHLNCVLRRLRLAREDDLEACASPP
 ALGGRYWAVGEFVCEPVVTHRPPPLAVPAPRALRCRAVGDPEPRVWSPQG
 RLLGNSRARAFNPTLLELVTEPFGDGIPTCIANAAGEATAVELTVGPPPPQLA
 NTSFCDPKPDALTPPSAASAKAVDTPPTDRGVTEHGNATAALVQVMDPDR
 IFLRMVQIQNSADLDLVRMIPAPRSRFLLDLARGTYDLCVLAIVEGSAATGLT
 ATPRVGCAEFSTEPALRCPGAPHPFLGGTMIITALGGVIVASVLFVFLLMRYKVHG
 QPPEKAKIPAPVSSVCSQINGALGPTTPAPPAPPEALRAHTVVRLLDCEPWGPGHE
 PVGP"

BASE COUNT 422 a 956 c 743 g 437 t
 ORIGIN
 Query Match 20.1%; Score 633; DB 6; Length 2558;
 Best Local Similarity 61.3%; Pred. No. 7.5e-100;
 Matches 1078; Conservative 0; Mismatches 665; Indels 16; Gaps 3;

QY	507	GTTTGGCGTGTGAGCGCTGCCCCCAAGTACTGTGTGTGCGCAGAACTGTCTGTGAGTCACT	566
DB	497	GTCAGCCACACCCAGCCCAATGTCCTGGCGGTGCGCGTCCAGACACAGTGTGCGCCCT	556
QY	567	GGGGACCTGTGCTCCAGGGCTGCTCTTTGACCCCTGATATGACCGCGGAC	626
DB	557	AAGCGTGTGTGCCAGGGGAGGCTCTGTTGTGTGCACTCCCTCGTGGACCCCGCGG	616
QY	627	AGTGGAGCTGCGCTGGCGGCAACTTATCATATCAATCAATCAATCAATCAATCAATCAAT	686
DB	617	AGCGGAGCTGCGGCTGGCAGACAACTTATCATGCTCCGCTCCGCTCCGCTCCGCTCCG	676
QY	687	CATGACGGGCTGGTGGACTGACCTGTCAGGACACATCAGCCACATCCAGCCCTT	746
DB	677	CATGACAGGCTGCTGCATCTGAGCTGTGCGGGAACACATCCGCGCAGTGGTCCGG	736
QY	747	TTCTTTCTGAGCTCGAGAGCTCCGCTCCCTGCACTTTGACAGAAATCGGCTCCCAAG	806
DB	737	CGCCTTCGCGACTGCGGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCG	796
QY	807	CCTTGGGGAGACACCTCCGCGGCTGFTCAACCTGACAGCACTTATCGTGAACAACAA	866
DB	797	ACTGGCGAGGGCAGCTGGCGGCTGTGTCACTTGGCCACTCATCTCAGAACAA	856
QY	867	CCAGCTGGGGCATCCAGATGAGGCTTTTGGAGACTTCTGCTGACATTTGGAGATCT	926
DB	857	CCAGCTGGCAGCGTGGCGGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCG	916
QY	927	GGACTCTCTTACAACAACTCCATGGCTGCGGACTCCGTCGGAGCACTGAGGATCTAA	986
DB	917	CGACTCTCTTACAACAACTCCATGGCTGCGGACTCCGTCGGAGCACTGAGGATCTAA	976
QY	987	CCTCACAGCTGAGCTGGACACCACTGCTGGATCACATCGCCGAGGGCACTTTTC	1046
DB	977	CGTCAACAGCTGGGCTTCGACCAACACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1036

QY	1047	AGACCTGAGAAAATGGCCCGCTGGATCTCACCTCAATCGGCTGCAGAAAGTGCCTCC	1106
DB	1037	CGCCTGCACAAGCTGGCCCGCTGGACATGACCTCCAAACCGCTGACCAACAATCCCA	1096
QY	1107	TGATCCCATTTGGCCCGCTCCAGGCTTCGGCTTTGACAGCCACACCTTTGGCCCA	1166
DB	1097	CGACCCACTTTCTCCCGCTGCTGCCAGGCCCCGGGCTCGCCGGCTCTGCG	1156
QY	1167	CTTGTCTTTAGTTTTGGGGTAAACCACTTACAGCAATTTGAGCTTCTGCTGCTGG	1226
DB	1157	CTTGTGTGGCTTTGGCGGAAACCCCTGACCTGCAACTGCGAAGCTGGTGTGGCTGG	1216
QY	1227	GAGCTCGAGCGGACGATGACTGGAAACCTGTGGCTCCCCAGGGGGCTCAAGGTTGG	1286
DB	1217	TCGCTTGGCGGGGAGGACGACCTGAGGCTTCCGCTGCGGCTGCTGGGGGGCGG	1276
QY	1287	CTACTTTGGCATGTGCTGAGGAGTTTGTGCGAGCCGCTCTCATCCCCAGCA	1346
DB	1277	CTACTTTGGCGGGGAGGAGGTTTTGTGCGAGCCCGCTGCTGACTCACC	1336
QY	1347	CACACAGTTTCTGTTCTGAGGCGCCAGGCGGCGCACACTCAAGTGCAGAAAGCA	1406
DB	1337	CTACCACTTCTGGCTGTGCCCGAGGTGGCGGCTGCGCTGCGCTGCGGGCAGTGG	1396
QY	1407	GGACCCAGCCCTTATCCACTGGGTAGCCCCCGATGACCCGCTGTTAGGAACTCTC	1466
DB	1397	GGACCCAGCCCTTATCCACTGGGTAGCCCCCGATGACCCGCTGTTAGGAACTCTC	1456
QY	1467	AAGGACCGTGTCTATGACAAATGGCACCTTGTGAGCTTCTCATCCACATCTCAG	1526
DB	1457	CCGTGCGCGGCTTCCCAATGGGACCTTGGAGTCTGCTCACCAGGCGGTTGATGG	1516
QY	1527	TGCTGCTTACCTGCAATGTGTCGAGAGGCGCCAGCCGCTGTTGGAGT	1586
DB	1517	TGGCATCTTACCTGCAATGGCGCAATGGAGTGGCGGCGCACAGCTGCTGTGGAG	1576
QY	1587	TTCCATGCTCAGCTGCCACACTCAGCAAGACAGCAGCCAGCCGCA--CTGACCC	1644
DB	1577	GACTGTGGTCCCCCACTCTCTCAGTAGCCACAGCACAGCTGTGACCCCGCG	1636
QY	1645	TCCCGCTCTCAGCATCACTGGTCCAGCAAGACAGCCGCGGAGGTGGAGGCA	1704
DB	1637	GGACGGGATCTGATGCTCTCACCCACCTCCGCTGCTCTTCTGCAAGGTGGC	1696
QY	1705	GGCGAGGCTCCAAAAGCCCCCGAAGCGGCTGCTTGTGTGAAAGTACCA	1764
DB	1697	CGACACTGGGCCCCCTA-----CCGACCGTGGCTCCAGTGTGACTGACACGGG	1748
QY	1765	ACCTGGGCTGGTCAAGTGTCTGTACAGCAAGTCCAGCAGCCCGGGTGAAGATG	1824
DB	1749	ACAGTGTCTTGTCCAGTGGCGGATCAGCGGCTTATCCCGGATCCGATGACCA	1808
QY	1825	CTCAGTACAATGCTCTGACATGAGTACTGATTTACAGGATGATCCAGCTCCAA	1884
DB	1809	ATCCAGTACAACAGCTCGGCTGATGATCCTCTGTTACAGGATGATCCCGGGAG	1868
QY	1885	AAGGCTTCTGTTCAACAACCTGTGTGAGGACTGCGTACGACTTGTGTGTGGCC	1944
DB	1869	CGTCTGCTTCTGTTGAGGACTGGGCTCAGGCGGACCTACGATGTTGCTGTGCTG	1928
QY	1945	ATGTGGGATGACACAGCAGCACACTCAGCGCCACCAACATCGTGGCTGCGCCAG	2004
DB	1929	GTGTATGAGGCGAGTGCACGGGCTCAAGGCGGCTGTTGGGCTGCGCCCGCTTC	1988
QY	2005	TTTCAACAGGCTGACTACCGCAGTGCAGTCCATGACAGCCAGATTTCTGGCGG	2064
DB	1989	TCCACCAACCTGCTGCGGCTATGCGGGCCATGCGGGCGCCGACGCTCCCTT	2048
QY	2065	ATGATCTGTTCTCGGGGCAATCATCGTGGCAGCGTGTGCTTCTCACTCGTCA	2124
DB	2049	ATGATCATCGGCTGGGCGGCTCATCGTAGCCTCGGTACTGTTCTTCTTCTGCTG	2108

ZymoGenetics, Inc. (US)
 Location/Qualifiers
 1. 1653
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 1. 1653
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="CAD23323.1"
 /db_xref="GI:18674973"
 /translation="MAPPLLLLLASGAACPLPVCVONLSESLTLCARHGLLFPVH
 NVDRVELRLADNFIQALGPPDFRMTGLVDLITLSENATRICARAFGLLESLSLH
 LDNRLVBLGTGSLRFPVNLQHLISLGNLQGLRIPAFDPDLFLESLEDLSYNNLRD
 PWAGIGAMPALHNLNHLNLDLPPGAFALQGLSRDLTSSRLAFLAIDLPLFSRGR
 DAEASPAVLVLSFSCNPLHRCCELLRLLRLLRLLPDDLETCPASPLAGRYFWAYPEGFF
 SCEPLLIARHTORLWLEGRATLRCALGDPAPFTMHWGPDRLVGNSSARAFPNAS
 TLEIGVTAGDAPGVTCTIATNPAGEATARVELRVLALPHGNSAEAGRPPSPDIA
 IPTASSHRHLKHLVPGADYDLCLLALSPPAGPSDLTATRLGLCAHFTSLPASPLCHA
 LQAHVLGGLTVAVGGVLAALTGLHCGPFWFGAGGPEMAASP"
 223 a 617 c 526 g 287 t

BASE COUNT
 ORIGIN
 Query Match 19.3%; Score 608; DB 6; Length 1653;
 Best Local Similarity 63.7%; Pred. No. 1.7e-95;
 Matches 1025; Conservative 0; Mismatches 540; Indels 43; Gaps 5;

520 GAGCCCTGCCCAAGTACTGTCTGCCAGAAATCTGTCTGAGTCACTGGGGACCCCTGTC 579
 QY
 43 GCGCCCTGCCCGTCCCTGCGCTGCCAGAACCTGTCCGAGTCTGCTGAGTCTGCT 102
 Db
 580 CCCTCCAAAGGGTCTCTTTGTACCCCTGATATTGACCGCGGACAGTGGAGTGCOC 639
 QY
 103 GCCCACCGAGCCCTGCTTTGTGCGCCCAACGCTGGACCGCGGCACAGTGGAGTGC 162
 Db
 640 CTGGGGCGCAACTTCATCCATCCAGCCCGCAGACTTTCGCAACATGACGGGGCTG 699
 QY
 163 CTGGTGAACCTTCATCCAGGCCCTGGGGCCCTGACTTCGCGAACATGACGGGACTG 222
 Db
 700 FTGGACCTGACCTGTCCAGGAACACCATCAGCCACATCCAGCCCTTTCTTTCTGGAC 759
 QY
 223 GTGGACCTGACACTGTCTCGCAATGCCATCACCGCATGGGGCCCGCCCTTTGGGGAC 282
 Db
 760 CTCGAGAGCTCCGCTCCCTGCATTTGACAGCAATCGGTGCCAAGCCCTGGGGAGGAC 819
 QY
 283 CTCGAGAGCTCCGCTCCCTGCATTTGACAGCAATCGGTGCCAAGCCCTGGGGAGGAC 342
 Db
 820 ACCCTCCGGGGCTGTCACCTGCAGCACCTTTCGTGAACAACAACACACAGCTGGGGCC 879
 QY
 343 AGCCTCCGGGGCCCGTCAATCTGCAGCACCTTTCGTGAACAACAACACACAGCTGGGGCC 402
 Db
 880 ATCGCAGATGAGCTTTGAGGACTTCTGCTGACATTTGGAGGATCTGGACCTCTCTAC 939
 QY
 403 ATCGCCCGGGAGCTTCCAGCCTTCTAGAGAGCTTGGAGGACTGGACCTGCTGCTTAC 462
 Db
 940 AACAACTCCATGGCTGCTGGGATCCGTCGACGCAATGCTCCACACTCCACACGCTG 999
 QY
 463 AACAACTCCGCGAGGTGCCCTGGGGCCGATCGGGCCATGCTGCTGCACACCTC 522
 Db
 1000 AGCTGGACCAACAACCTGCTGGATCACATCGCCGAGGGGACCTTTGTCAGACCTGCGAGAA 1059
 QY
 523 AACCTGACCAACAACCTTATGACGCACTGCCCCAGGGCCCTTCGCCAGCTCGGTGAG 582
 Db
 1060 CTGGCCCGCTGATCTCCTCAATCGGTGACAGAACTGCCCCCTGATCCCATCTTT 1119
 QY
 583 CTCTCCCGCTGGACCTCACCTCAACCGCTGGCCACCTGGCTCGGACCCGCTTTTC 642
 Db
 1120 GCCCGCTCCAGGCTTCGGCTTTGACAGCCACACCTTTGCCACCCCTTGTCTTCTT 1179
 QY
 643 TCTCGTGGCGTGA-----GCAGAGCCCTCTCCCGCCCTCTGGTGGCGGCTCGAGCG 690
 Db
 1180 TTTGGGGTAAACCCACTTCTGCAATGTGAGCTTCTGGTGGCGGAGGCTCGAGCGG 1239
 QY

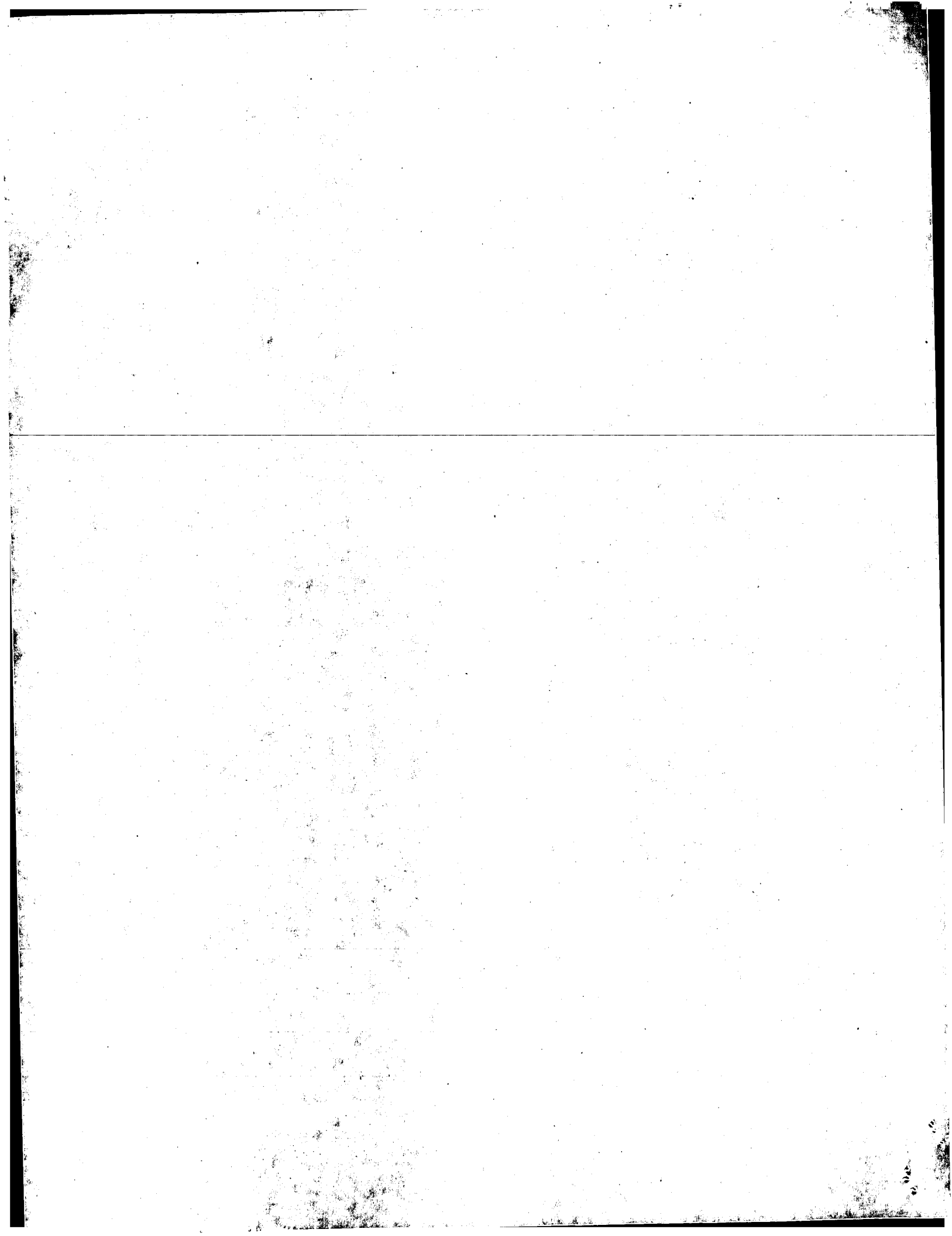
1480 TATGCAATGGACCTTGACATCTTATACACATCTCAGACAGTGTGCTTACC 1539
 Db
 1243 TTTCCCAACCGGACCTTAGATTTGGGTGACCGCGCTGGGACCGTGGGGCTACAC 1302
 QY
 1540 TGCATTGCTGCAATGCTGCGGAGAGCCGACGTCATGCTGCTGCTGCTGCTGCTG 1599
 Db
 1303 TGCATCGCCACCAACCTTCTGTTGAGCCACAGCCGAGTAGACTCGGGTCTGCGCC 1362
 QY
 1600 CTGCCACACTCAGCAACAGCAGCCAGCCGACTGACCCCCCAAGTCCCGCTCTCAGAC 1659
 Db
 1363 TTGCCCCATGTTGGAAACAGCACTGCCGAGGGGGCCG-----CGGGCCCTCGGAC 1416
 QY
 1660 ATCACTGGTCTCCAGCAAGACAGCCGGGGAGTGGAGCAGTGGGGCGGAGAGCTCC 1719
 Db
 1417 ATCGCGCTCCCTCGCACTGCTGCGAGGTTGAGGGACCTCGAGTCTGAGC----- 1472
 QY
 1720 AAMAGCCCCCGGACCGGGCTGCTGTTGCTCTGAAGTACCCACCACTCGGCGCTGCTG 1779
 Db
 1473 -----AGCGTGCAGTGCAGGAGTGCAGGAGTGCAGCCACCTCAGGGTGGT 1515
 QY
 1780 AAGTGGTCTCTCAGCAAGTGCAGACCCCGGGTGAAGTGTACAGTGCAGTACAACCTGC 1839
 Db
 1516 AGTGGGTCCCGGGCCGACCCAGTGTGGATGTTCCAAATCCAGTACACACAGC 1575
 QY
 1840 TCTGACGTAGGTACTGATTTACAGATGATCCAGCTCCCAACAGGCTTCTGTTGCTG 1899
 Db
 1576 AGCGAAGATGAGACCTCACTACCGGATTTCTCCAGCCCTCAGCCACCACTTCTGCTG 1635
 QY
 1900 AACAACTCGTTCAGGACTGGTACGACTTTGTGTGTGTTGTTGTTGTTGTTGTTGTTG 1959
 Db
 1636 AAGCACTCGTCCCGGGCTGACTATGACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1695
 QY
 1960 G---CCACGACTCAGCCGACCAACATCTGGGGCTGGCCAGTTTTCACCAAGGCT 2016
 Db
 1696 GGGCCCTTGCCTCAGGCGCACAGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGT 1755
 QY
 2017 GACTACCCAGTGCAGTCCATGACAGCCGATTTCTGGGGCCAGCCATGCTGCTGCTG 2076
 Db
 1756 GCCTGCCCCCTGTCAGCCCTGACGGCCACGCTGTGGGGGACCCCTGACCGTGGCC 1815
 QY
 2077 ATCGGGGACATCTGTCGACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2136
 Db
 1816 GTGGGGGTGTGCTGTTGCTTACTGTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1875
 QY
 2137 AAGTCTGCAACACAGGCCCCCAGCAAGATGCGAGGCGCTGAGCAATGTGACTCG 2196
 Db
 1876 CGGGGGCGGAAATGGCCCTTCCCTCAAG-----CTCAGCCAGTCCAGTCC 1926
 QY
 2197 CAGACCAAGGGCCGACCGCTCCAGGAGGCGCACCGGGGGCCCGCCCGGAG 2256
 Db
 1927 CAGACCAAGGGCCCGGACCGCCACACCCAGGCCACCGCCCGGGAGCCCCCGGCC 1986
 QY
 2257 GCGCCGCG 2265
 Db
 1987 CGGCGCGAG 1995

RESULT 15
 AX358336 1653 bp DNA linear PAT 13-FEB-2002
 LOCUS
 DEFINITION Sequence 7 from Patent WO202604.
 ACCESSION AX358336
 VERSION AX358336.1 GI:18674972
 KEYWORDS human.
 SOURCE Homo sapiens
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
 AUTHORS Thayer, F. C., Sheppard, P. O. and Presnell, S. R.
 TITLE Leucine-rich repeat proteins, zlr17, zlr18 and zlr19
 JOURNAL Patent: WO 0202604-A 7 10-JAN-2002;

Db 691 TTTAGGGGAACCCCTGCACTGCAACTGTGAGTCTGTGGCTGGCGGGCTGGCGGG 750
 Qy 1240 GACGATGACCTGAAACCTGTGGCTCCCGAGGGGGCTCAAGGGTGGCTACTTCTGGCAT 1299
 Db 751 CCGGAGGACCTGAAACGTCGGCTCCCGCCGGCTGGCGGGCTACTTCTGGCA 810
 Qy 1300 GTGGTGTGAGGAGTGTGTGCGAGCCGCTCTCATCCAGCAGCACAACAAAGTTG 1359
 Db 811 GTGCCGAGGGGAGTCTCTGTGAGCGCCCTCATTTGCCCGCCACACGAGCGCTC 870
 Qy 1360 CTGGTCTGAGGGCCAGGGCCACACTCAAGTGAAGCCATTTGGGGACCCAGCC 1419
 Db 871 TGGGTGCTGGAAGCCAGCGGGCCACGCTGGGTGGGGCCCTGGGTGACCCCGGGCT 930
 Qy 1420 CTTATCCACTGGGTAGCCCGGATGACCGCTGGTAGGAACTCCTCAAGGACCGGTGC 1479
 Db 931 ACCATGCACTGGGTGGTCTGACGACCGGTTGGTTGGCACTCCTCCCGAGCCGGGCT 990
 Qy 1480 TATGACAATGGACCCCTGGACATCTTATCACCACATCTCAGGACAGTGGTCCCTTACC 1539
 Db 991 TTCCCAACCGGACCTTAGAGATTGGGGTACCCGGCTGGGACGCTGGGGCTACACC 1050
 Qy 1540 TGCATTGTGCCAATGCTCCGAGAGCCAGCCATGGTGGAGGTCTCCATCGTCCAG 1599
 Db 1051 TGCATCGCCACCAACCTGCTGTGAGGCCACAGCCCGAGTAGAACTGCGGGTGTGGCC 1110
 Qy 1600 CTGCCACACTCAGCAACAGCACCGCCGCACTGCACCCCCAAAGTCCCGCTCTCAGAC 1659
 Db 1111 TTGCCCATGGTGGBAACAGAGTGCAGGGGGGGCCG-----CCGGCCCTCGGAC 1164
 Qy 1660 ATCACTGGTTCAGCAAGACACAGCCGGGAGTGGAGGAGTGGGGGGGAGAGCTCCC 1719
 Db 1165 ATGCCCGCTCCGCTCGCACTGTGCGAGGTTGAGGGACGCTGGAGTCTGAGCC---- 1220
 Qy 1720 AAAAGCCCCCGAAGCGGCTGCTTGTCTGTAAGTACCACACACTCGGCCCTGGTGC 1779
 Db 1221 -----AGCCGTGCAGGTGACGGAGGTGACCCCACTCAGGGCTGGGTG 1263
 Qy 1780 AAGTGTCTGTGCAAGTCAAGTCCAGCCCGGGTGAAGATGACCGCTGCACTCAACTGC 1839
 Db 1264 AGCTGGGGTCCCGGGCCAGCCGACCCAGTGTGATGTTCCAAATCCAGTACAACAGC 1323
 Qy 1840 TGTGAGATGAGTACTGATTTACAGGATGATCCAGCCCTCCAAAGGCTTCTGGTGC 1899
 Db 1324 AGCGAAGATGAGACCTCATCTACCGGATTTCCAGCCCTCCAGCCCACTTCTGCTG 1383
 Qy 1900 AACAACTGGTGTGAGGACTGCTAGGACTTGTGTGCTGGCCATGGGATGACACA 1959
 Db 1384 AAGCACTCGTCCCGGGCTGACTATGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1443
 Qy 1960 G---CCACGACTCAGGCCCAACAATCGTGGGCTGGCCAGTTCACCAAGGCT 2016
 Db 1444 GGGCCCTGTGACCTCAGGCCACAGGCTGTGGGTGTGCCATTTCTCCAGCTGGCG 1503
 Qy 2017 GACTACCGCAGTGCAGTCCATGCAAGCCAGATTTCTGGGGCCACCAATGATTCCTGGTC 2076
 Db 1504 GCCTCGCCCTGTGCCACCGCCCTGCAGGCCACGCTGTGGGGGGACCCCTGACCCGTGGCC 1563
 Qy 2077 ATCGGGGATCATCGTGGCAGCC-TGCTGGTCTTCATCGTCACT 2123
 Db 1564 GTGGGGGTGTGCTGGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1611

Search completed: June 8, 2003, 18:21:50
 Job time : 9339 secs



GenCore version 5.1.i.6
 Copyright (c) 1993 - 2003 CompuGen Ltd.
 OM nucleic - nucleic search, using sw model
 Run on: June 8, 2003, 15:09:17 ; Search time 664 Seconds
 (without alignments)
 10663.076 Million cell updates/sec

Title: US-09-831-846-1
 Perfect score: 3144
 Sequence: 1 gctgtgctccctcctcgtga.....gctgttggattgcttatg 3144
 Scoring table: IDENTITY NUC
 Gapop 10.0 , Gapext 1.0
 Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478
 Minimum DB seq length: 0
 Maximum DB seq length: 2000000000
 Post-processing: Minimum Match 0%
 Maximum Match 100%
 Listing first 45 summaries

Database : N_Geneseq_101002.*
 1: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
 2: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
 3: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
 4: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
 5: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
 6: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
 7: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
 8: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
 9: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
 10: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
 11: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
 12: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
 13: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
 14: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
 15: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
 16: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
 17: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
 18: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
 19: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
 20: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
 21: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
 22: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
 23: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
 24: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	3144	100.0	3144	21 AAA40083	Human brain-specific
2	3144	100.0	3144	21 AAA60605	Human brain-specific
3	2780.2	88.4	2818	22 AA544992	CDNA encoding nove
4	2780.2	88.4	2818	22 AA158215	Human polynucleoti
5	969	30.8	969	23 AA577723	DNA encoding nove
6	678.8	21.6	771	22 AA160001	Human polynucleoti
7	634.6	20.2	1887	22 AAH25568	Nucleotide sequenc
8	634.6	20.2	1887	22 AAH78215	Nucleotide sequenc
9	633	20.1	1887	22 AA166986	Human LRR family m

10	633	20.1	2637	22	AAI66985	Human leucine-rich
11	622.6	19.8	2369	24	AAAD38695	Human LP220 secret
12	621.8	19.8	2855	24	ABK70010	CDNA encoding huma
13	618.8	19.7	2316	22	AAH78204	Nucleotide sequenc
14	608	19.3	1653	24	AAAD28124	Human leucine-rich
15	554.6	17.6	1653	24	AAAD28125	Human leucine-rich
16	546	17.4	3012	24	ABN93416	Human secreted pro
17	544.4	17.3	2301	22	AAH78197	Nucleotide sequenc
18	477.2	15.2	1440	22	AAH78214	Nucleotide sequenc
19	477.2	15.2	8269	22	AAI99642	Human expressed po
20	477.2	15.2	8269	22	ABA06793	Human genomic DNA
21	477.2	15.2	8269	22	AAK90941	Human digestive sy
22	477.2	15.2	8269	22	AAI64010	Human polynucleoti
23	477.2	15.2	8269	22	AAAS3472	Human DNA for a no
24	477.2	15.2	8269	22	AAAS31976	Human liver associ
25	477.2	15.2	8269	22	AAAS33436	DNA encoding human
26	477.2	15.2	8269	22	AAAS35069	DNA #19 encoding h
27	477.2	15.2	8269	24	ABQ66796	Human polynucleoti
28	477.2	15.2	8269	24	ABN90331	Human liver antige
29	471	15.0	1510	23	AAAS70133	DNA encoding nove
30	471	15.0	1602	22	AAH78203	Nucleotide sequenc
31	462.2	14.7	1582	23	AAAS69014	DNA encoding nove
32	461.4	14.7	1872	22	AAF76853	Human secreted pro
33	396	12.6	441	22	AAAS45180	CDNA encoding nove
34	298.2	9.5	1268	22	AAAS29522	Human endocrine po
35	226.2	7.2	541	22	AAI99624	Human expressed po
36	226.2	7.2	541	22	AAAS33219	DNA encoding human
37	204.2	6.5	432	24	ABK33084	DNA encoding nove
38	204.2	6.5	432	24	AAAD28122	Human leucine-rich
39	194	6.2	902	22	ABA06401	Human cDNA SEQ ID
40	194	6.2	902	22	AAAS29645	Human endocrine po
41	194	6.2	902	22	AAI63902	Human polynucleoti
42	194	6.2	902	22	AAAS31239	Human cDNA encodin
43	194	6.2	902	22	AAAS34784	CDNA encoding nove
44	194	6.2	902	24	ABQ66563	Human polynucleoti
45	193.2	6.1	902	24	ABL90356	Human polynucleoti

ALIGNMENTS

RESULT 1
 AAA40083
 ID AAA40083 standard; cDNA; 3144 BP.
 XX
 AC AAA40083;
 DT 19-OCT-2000 (first entry)

Human brain-specific transmembrane glycoprotein encoding cDNA.
 DE
 DE Transmembrane glycoprotein; human; brain; cellular signal transducer;
 KW neuroactive; neuroprotective; cerebroprotective; drug development;
 KW treatment; nervous disease; diagnostic; ss.
 XX
 OS Homo sapiens.
 XX
 XX Location/Qualifiers
 FH 466..2835
 FT /*tag= a
 FT /product= "Transmembrane glycoprotein"
 XX
 XX WO200031256-A1.
 XX
 XX PD 02-JUN-2000.
 XX
 XX PF 18-NOV-1999;. 99WO-JP06449.
 XX
 XX PR 20-NOV-1998; 98JP-0331727.
 XX
 XX PA (CHUG-) CHUGAI RES INST MOLECULAR MEDICINE INC.
 XX (KAZU-) KAZUSA DNA RES INST.

Funahashi S, Miyata S, Nomura N, Nagase T, Ohara O;

WPI: 2000-411954/35.
P-PSDB; AAB09986.

Gene encoding a brain-specific transmembrane glycoprotein that has a typical PDZ protein binding motif and functions as a cellular signal transducer, useful in developing drugs for treating nervous diseases - Claim 1b; Page 43-51; 63pp; Japanese.

This invention describes a novel gene encoding a human brain-specific transmembrane glycoprotein that has a typical PDZ protein binding motif and functions as a cellular signal transducer. The product of the invention has neuroactive, neuroprotective and cerebroprotective activity. The gene and encoded protein are useful in developing drugs for treating nervous diseases, and also for studying functions of the nervous system or onset mechanism of nerve-related diseases. They may also be used for the production of diagnostic reagents such as oligonucleotide probes and antibodies for detecting proteins comprising PDZ domains and the nucleic acids that encode them. They may also be used to produce and identify modulators of the proteins expression and activity such as antisense sequences and antibodies. This sequence encodes the human brain-specific transmembrane glycoprotein described in the method of the invention.

Sequence 3144 BP; 589 A; 1027 C; 919 G; 609 T; 0 other;

Query Match 100.0%; Score 3144; DB 21; Length 3144;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3144; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Table with 4 columns: Line number, Query sequence, Subject sequence, and Match status. It shows a 100% identity match between the query and subject sequences across all 600 positions.

Table with 4 columns: Line number, Subject sequence, Match status, and Line number. It shows a 100% identity match between the query and subject sequences across all 600 positions.

Db 1681 ACCCGGGAGGTGGAGGAGTGGGGCGGAGAGCCTCCAAAAGCCCCCCCCGCGGT 1740
 Qy 1741 GTGTTGTGTGAAGTACCACACCTCGCCCTGGTCAAGTGTGTGTGACAGATCA 1800
 Db 1741 GTGTTGTGTGAAGTACCACACCTCGCCCTGGTCAAGTGTGTGTGACAGATCA 1800
 Qy 1801 GCACCCGGGTGAAGATGACAGTGCAGTCAACTGCTGACAGTGAAGTACTGATT 1860
 Db 1801 GCACCCGGGTGAAGATGACAGTGCAGTCAACTGCTGACAGTGAAGTACTGATT 1860
 Qy 1861 TACAGATGATCCAGCCTCCAAAGGCTTCTGGTCAACAACCTGGTGTGAGGACT 1920
 Db 1861 TACAGATGATCCAGCCTCCAAAGGCTTCTGGTCAACAACCTGGTGTGAGGACT 1920
 Qy 1921 GCTACGACTTGTGTGTGCTGCCATGTGGATGACACAGCCAGCACTCACGGCCACC 1980
 Db 1921 GCTACGACTTGTGTGTGCTGCCATGTGGATGACACAGCCAGCACTCACGGCCACC 1980
 Qy 1981 AACATCGTGGTGGCCAGTCTTCCAAAGGCTTCTGGTCAACAACCTGGTGTGAGGACT 2040
 Db 1981 AACATCGTGGTGGCCAGTCTTCCAAAGGCTTCTGGTCAACAACCTGGTGTGAGGACT 2040
 Qy 2041 CACAGCAGATCTGGCGGACCAATGATCTGGTCACTCGGGGATCATCTGGCCAGG 2100
 Db 2041 CACAGCAGATCTGGCGGACCAATGATCTGGTCACTCGGGGATCATCTGGCCAGG 2100
 Qy 2101 CTGCTGCTTCATCGTCACTCATGCTGCTCAAGGCTGCTCAAGCCACAGGCCCC 2160
 Db 2101 CTGCTGCTTCATCGTCACTCATGCTGCTCAAGGCTGCTCAAGCCACAGGCCCC 2160
 Qy 2161 ACCAAGTGGCAGCGCCGTGAGCAATGTTACTGCGAGCAACAAGCCGCCACGACC 2220
 Db 2161 ACCAAGTGGCAGCGCCGTGAGCAATGTTACTGCGAGCAACAAGCCGCCACGACC 2220
 Qy 2221 CCTCAAGCAGCGCACCAGCCGGCCCGCCGCGAGGCCCCCGCAAGTGTGGTGGC 2280
 Db 2221 CCTCAAGCAGCGCACCAGCCGGCCCGCCGCGAGGCCCCCGCAAGTGTGGTGGC 2280
 Qy 2281 AACGAGCTCTGGACTTCAACCGCCAGCTGGCCGCGGAGTCTCTCTCCCTCCAGC 2340
 Db 2281 AACGAGCTCTGGACTTCAACCGCCAGCTGGCCGCGGAGTCTCTCTCCCTCCAGC 2340
 Qy 2341 TCCCTGGGAGTGGGAGGCTCGGGGCTGGAGCCGCCCCCTGGAGGATCCACCTCC 2400
 Db 2341 TCCCTGGGAGTGGGAGGCTCGGGGCTGGAGCCGCCCCCTGGAGGATCCACCTCC 2400
 Qy 2401 GCCCGGCCCCAAGCCAGCTTCAACCGCCAGCTGGCCGCGGAGTCTCTCTCCCTCCAGC 2460
 Db 2401 GCCCGGCCCCAAGCCAGCTTCAACCGCCAGCTGGCCGCGGAGTCTCTCTCCCTCCAGC 2460
 Qy 2461 AAGAGTCAGAGAAAGAGGAGTGTGGACTCCAGGACTCCAGCCGGGAGAGGGGTGGG 2520
 Db 2461 AAGAGTCAGAGAAAGAGGAGTGTGGACTCCAGGACTCCAGCCGGGAGAGGGGTGGG 2520
 Qy 2521 AGTGGCCCGGGCCACACTCGGACCGAGACCACTGCTGGGGCCCCCTGGGGCCCGG 2580
 Db 2521 AGTGGCCCGGGCCACACTCGGACCGAGACCACTGCTGGGGCCCCCTGGGGCCCGG 2580
 Qy 2581 GCCAGGAGCTGCTCCCTTCCCTGGAGGCGCAAGCCAGCCAGCCACTCTCCGAC 2640
 Db 2581 GCCAGGAGCTGCTCCCTTCCCTGGAGGCGCAAGCCAGCCAGCCACTCTCCGAC 2640
 Qy 2641 ATGGGGACTTTGCTGCTCGGGCGGGAGGCTGCTCGGGGGCTTACAGTCTCTCT 2700
 Db 2641 ATGGGGACTTTGCTGCTCGGGCGGGAGGCTGCTCGGGGGCTTACAGTCTCTCT 2700
 Qy 2701 CGGAAGGTCTGAAACATCTGGAACGAGCCGCTCTCTGTCAAAGGCAATGCTTTGCC 2760
 Db 2701 CGGAAGGTCTGAAACATCTGGAACGAGCCGCTCTCTGTCAAAGGCAATGCTTTGCC 2760
 Qy 2761 TTTGAGGAGTACTGTTGGGGCCCCGGGACTTTTGGCAGCTCCGAATGGGTGATG 2820
 Db 2761 TTTGAGGAGTACTGTTGGGGCCCCGGGACTTTTGGCAGCTCCGAATGGGTGATG 2820

Qy 2821 GAGACACGGTCTAGGTGGGGTGGCATGCTCCCTTCTGTGCGCAGGGTGGAGAAG 2880
 Db 2821 GAGACACGGTCTAGGTGGGGTGGCATGCTCCCTTCTGTGCGCAGGGTGGAGAAG 2880
 Qy 2881 GGGAAAGATCTCACATGCGCAGTGTGTTGGAGTTTCCATGCTGATGTTACATCCAGG 2940
 Db 2881 GGGAAAGATCTCACATGCGCAGTGTGTTGGAGTTTCCATGCTGATGTTACATCCAGG 2940
 Qy 2941 ACAGTTCCTCCCTGTCATGCGCTGCTGTCCTCCCTCCCTCCCTCCCTCCATCA 3000
 Db 2941 ACAGTTCCTCCCTGTCATGCGCTGCTGTCCTCCCTCCCTCCCTCCCTCCATCA 3000
 Qy 3001 CCTCCCAACCCCGCGGGTGTGCTCAGGAAATGTGACTCGTCAAAATGCCGGACT 3060
 Db 3001 CCTCCCAACCCCGCGGGTGTGCTCAGGAAATGTGACTCGTCAAAATGCCGGACT 3060
 Qy 3061 GAGCCCTGAGTGTGTTGAAAGCGGAGACTCCCGCTTCTAATCACAATGTAGCTTCAA 3120
 Db 3061 GAGCCCTGAGTGTGTTGAAAGCGGAGACTCCCGCTTCTAATCACAATGTAGCTTCAA 3120
 Qy 3121 GCAAGCGGCTTGGATTGCTTATG 3144
 Db 3121 GCAAGCGGCTTGGATTGCTTATG 3144

RESULT 2
 AAA60605
 ID AAA60605 standard; cDNA; 3144 BP.
 XX AAA60605;
 AC
 DT 19-OCT-2000 (first entry)
 XX
 DE Human hh00149 protein encoding cDNA SEQ ID NO:3.
 XX
 KW Human; ubiquitin-like protein; l49y2H#151; hh00149; brain;
 2-hybrid screening; neuroprotective; signal transducer;
 nervous system disease; diagnosis; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 466..2835
 FT /*tag= a
 FT /product= "hh00149"
 XX
 PN WO200031255-A1.
 XX
 PD 02-JUN-2000.
 XX
 PF 18-NOV-1999; 99WO-JP06448.
 XX
 PR 20-NOV-1998; 98JP-0331701.
 XX
 PA (CHUG-) CHUGAI RES INST MOLECULAR MEDICINE INC.
 XX
 PI Funahashi S, Miyata S;
 XX
 XX WPI; 2000-400066/34.
 DR P-PSDB; AAB12448.
 XX
 PT Gene encoding ubiquitin-like protein which interacts with protein
 hh00149, useful in the diagnosis and treatment of diseases associated
 with the nervous system.
 XX
 PS Example 2; Page 64-73; 88pp; Japanese.
 XX
 CC The present invention describes a ubiquitin-like protein, designated
 CC l49y2H#151, which interacts with protein hh00149 expressed specifically
 in the brain. The l49y2H#151 protein has neuroprotective activity, and
 CC is a signal transducer. The l49y2H#151 gene and encoded protein are
 CC useful in the diagnosis and treatment of diseases associated with the

2041 CACAGCCAGATTCGGGGGACCAATGATCCTGGTTCATCGGGGGCATCATCTGTGGCCAGG 2100
 Db |||||
 2041 CACAGCCAGATTCGGGGGACCAATGATCCTGGTTCATCGGGGGCATCATCTGTGGCCAGG 2100
 Qy |||||
 2101 CTGCTGGTTCATCGTTCATCTCTGTCAGAGTCTCAAGTCTGCAACACGAGCCGCC 2160
 Db |||||
 2101 CTGCTGGTTCATCGTTCATCTCTGTCAGAGTCTCAAGTCTGCAACACGAGCCGCC 2160
 Qy |||||
 2161 AGCAAGATGGCAGCGCCGTGAGCAATGTGTACTGCGAGACCAAGGGCCCGCCAGCCCG 2220
 Db |||||
 2161 AGCAAGATGGCAGCGCCGTGAGCAATGTGTACTGCGAGACCAAGGGCCCGCCAGCCCG 2220
 Qy |||||
 2221 CCTCAAGAGCGCACCGCCGGCCCGCCGAGGCGCCCGCGAAGTGTGTGTGCGC 2280
 Db |||||
 2221 CCTCAAGAGCGCACCGCCGGCCCGCCGAGGCGCCCGCGAAGTGTGTGTGCGC 2280
 Qy |||||
 2281 AACGAGCTCCTGACTTACCCGCCAGCTGGCCCGCCGAGTCTCTTCCCTCCAGC 2340
 Db |||||
 2281 AACGAGCTCCTGACTTACCCGCCAGCTGGCCCGCCGAGTCTCTTCCCTCCAGC 2340
 Qy |||||
 2341 TCCCTGGGAGTGGGAGGCTCGGGGCTGGGACCGGCCCTTGGAGATCCACCTCC 2400
 Db |||||
 2341 TCCCTGGGAGTGGGAGGCTCGGGGCTGGGACCGGCCCTTGGAGATCCACCTCC 2400
 Qy |||||
 2401 GCCCGCGCCCAAGCCAGCCTTGACCCCTGATGGGGCCCTTCCCTCCCTGGACCTC 2460
 Db |||||
 2401 GCCCGCGCCCAAGCCAGCCTTGACCCCTGATGGGGCCCTTCCCTCCCTGGACCTC 2460
 Qy |||||
 2461 AAGAGTCAGAGAAAGGAGAGTGTGGACTCCAGGACTCCAGCCGGGAGAGGGGCTGG 2520
 Db |||||
 2461 AAGAGTCAGAGAAAGGAGAGTGTGGACTCCAGGACTCCAGCCGGGAGAGGGGCTGG 2520
 Qy |||||
 2521 ACCTCGGGCCGGCCACCACTCGGACCGAGACCACTCTCGGGGCCCCCTCGGGCCCGG 2580
 Db |||||
 2521 ACCTCGGGCCGGCCACCACTCGGACCGAGACCACTCTCGGGGCCCCCTCGGGCCCGG 2580
 Qy |||||
 2581 GCCAGAGCTCCTCCCTTCCCTTGGAGGCAAGCCGCAAAACGAGCCACTCTTTCGAC 2640
 Db |||||
 2581 GCCAGAGCTCCTCCCTTCCCTTGGAGGCAAGCCGCAAAACGAGCCACTCTTTCGAC 2640
 Qy |||||
 2641 ATGGGGACTTTGCTGCTCGGGCGGGAGGGTCTGTCGGGGGCTACAGTCTCTCT 2700
 Db |||||
 2641 ATGGGGACTTTGCTGCTCGGGCGGGAGGGTCTGTCGGGGGCTACAGTCTCTCTCT 2700
 Qy |||||
 2701 CGGAAGGTTCGAACAATCGAGCCAGCCTCTCTGTCAACGGCATGTCTTTGCC 2760
 Db |||||
 2701 CGGAAGGTTCGAACAATCGAGCCAGCCTCTCTGTCAACGGCATGTCTTTGCC 2760
 Qy |||||
 2761 TTTGAGGAGTGACTGTGGGGCCCGGGGACTTTTGGCAGTCCGAAATGGGTGATG 2820
 Db |||||
 2761 TTTGAGGAGTGACTGTGGGGCCCGGGGACTTTTGGCAGTCCGAAATGGGTGATG 2820
 Qy |||||
 2821 GAGGACCGTCTAGTGGGGTGGGCGATGCTCCCTTTCTGTCGGCAGGGTGGGAGAG 2880
 Db |||||
 2821 GAGGACCGTCTAGTGGGGTGGGCGATGCTCCCTTTCTGTCGGCAGGGTGGGAGAG 2880
 Qy |||||
 2881 GGGAAAGATTCACATGGCAAGTGTGTGGAGTGTTCATGTTTACATCCAGGG 2940
 Db |||||
 2881 GGGAAAGATTCACATGGCAAGTGTGTGGAGTGTTCATGTTTACATCCAGGG 2940
 Qy |||||
 2941 ACAGTTTCTCTCCCTGTCGAAAGGCGAGCTCCGCTTCTTAATCACAATAATGAGCTACAA 3000
 Db |||||
 2941 ACAGTTTCTCTCCCTGTCGAAAGGCGAGCTCCGCTTCTTAATCACAATAATGAGCTACAA 3000
 Qy |||||
 3001 CCTCCCACACCCCGGGGTGTCTCAGGAAATGTGGACTCGTCAAAATCCCGGACT 3060
 Db |||||
 3001 CCTCCCACACCCCGGGGTGTCTCAGGAAATGTGGACTCGTCAAAATCCCGGACT 3060
 Qy |||||
 3061 GAGCCCTGAGTGTGGAAAGGCGAGCTCCGCTTCTTAATCACAATAATGAGCTACAA 3120
 Db |||||
 3061 GAGCCCTGAGTGTGGAAAGGCGAGCTCCGCTTCTTAATCACAATAATGAGCTACAA 3120
 Qy |||||
 3121 GCAAGCGGCTTTGGATTGCTTATG 3144

Db 3121 GCAAGCGGCTTTGGATTGCTTATG 3144
 |||||
 RESULT 3
 AAS44992
 ID AAS44992 standard; cDNA; 2818 BP.
 XX AAS44992;
 XX AC
 XX DT 18-DEC-2001 (first entry)
 XX DE cDNA encoding novel human secretory protein, Seq ID No 73.
 XX KW Human; secreted protein; arthritis; Crohn's disease; sepsis; shock;
 KW ischaemia-reperfusion injury; haematopoiesis; cancer; neuropathy;
 KW transgenic animal; Alzheimer's disease; Parkinson's disease; burn;
 KW amyotrophic lateral sclerosis; platelet disorder; thrombocytopenia;
 KW ulcer; osteoporosis; bone degenerative disorder; periodontal disease;
 KW gut protection; lung; liver fibrosis; immune deficiency; infection;
 KW severe combined immunodeficiency; SCID; autoimmune disorder; allergy;
 KW multiple sclerosis; rheumatoid arthritis; diabetes mellitus; asthma;
 KW fertility; analgesic; pain; antigen; ss.
 XX OS Homo sapiens.
 OS WO200166689-A2.
 XX PN
 XX PD 13-SEP-2001.
 XX PF 05-MAR-2001; 2001WO-US04942.
 XX PR 07-MAR-2000; 2000US-0519705.
 PR 19-MAY-2000; 2000US-0574454.
 PR 17-JUN-2000; 2000US-0596193.
 PR 14-JUL-2000; 2000US-0616847.
 PR 19-SEP-2000; 2000US-0665363.
 PR 20-OCT-2000; 2000US-0693267.
 XX PA (HYSE-) HYSEQ INC.
 XX PI Tang YT, Liu C, Asundi V, Xu C, Wehrman T, Ren F, Ma Y, Zhou P;
 PI Zhao QA, Yang Y, Drmanac RT, Zhang J, Chen R, Xue AJ, Wang J;
 XX WI 2001-589934/66.
 DR P-PSDB; AAU28092.
 XX PT Novel polypeptides and nucleic acids obtained from cDNA libraries
 prepared from various human tissues, for diagnosis and treatment of
 cancer, neurological, inflammatory, and autoimmune disorders.
 XX PS Claim 1; SEQ ID No 73; 107pp; English.
 XX CC The invention relates to novel isolated human secreted polypeptides (I)
 and polynucleotides (II). (I) and (II) are useful for treating
 CC inflammatory conditions such as arthritis, nephritis, Crohn's disease,
 CC ischaemia-reperfusion injury, shock, sepsis, immune responses, and is
 CC involved in increasing haematopoiesis, stem cell survival, bone growth
 CC and remodeling. (I), (II) and modulators of (II) are useful for
 CC prophylaxis or treatment of one or more cancers. (II) is also useful for
 CC creating transgenic animals useful for studying the in vivo activities of
 CC the polypeptide as well as for studying modulators of the polypeptides.
 CC (I) induces the proliferation of neural cells and regeneration of nerve
 CC and brain tissue and is useful for the treatment of central and
 CC peripheral nervous system diseases and neuropathies, such as Alzheimer's,
 CC Parkinson's disease, Huntington's disease, and amyotrophic lateral
 CC sclerosis. In addition, (I) is involved in chemotactic or chemokinetic
 CC activity, regulation of haematopoiesis and is useful for treating myeloid
 CC or lymphoid cell disorders, platelet disorders such as thrombocytopenia
 CC and for regeneration of bone, cartilage tendon, ligament and/or nerve
 CC tissue growth, and in tissue repair, healing of burns, incisions,
 CC ulcers, for treating osteoporosis, osteoarthritis, bone degenerative
 CC disorders, or periodontal disease. Furthermore, (I) is also useful for

gut protection or regeneration and treatment of lung or liver fibrosis, reperfusion injury in various tissues, various immune deficiencies and disorders including severe combined immunodeficiency (SCID), bacterial or fungal infections, autoimmune disorders e.g. multiple sclerosis, rheumatoid arthritis, diabetes mellitus, myasthenia gravis, allergic reactions and conditions, such as asthma or other respiratory problems. In addition, (I) affects biorhythms or circadian cycles of rhythms, fertility, metabolism, catabolism, anabolism, storage or elimination of dietary fat, lipid, protein, carbohydrate, vitamins, minerals, provides analgesic effects or other pain reducing effects, immunoglobulin like activity and can act as an antigen in a vaccine composition to raise an immune response. AAS44920-AAS45295 represent novel human secreted protein coding sequences of the invention.

XX Sequence 2818 BP; 523 A; 933 C; 825 G; 537 T; 0 other;
 Query Match 88.4%; Score 2780.2; DB 22; Length 2818;
 Best Local Similarity 99.7%; Pred. No. 0;
 Matches 2785; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy	43	CACAAACCGCGAGCTCGTGAACCTTGAAGACACCCACATTCACAAAGATGCCGAGGTT	102
Db	26	CCCCCCCCCTGGCTCGTCTGAACCTTGAAGACACCCACATTCACAAAGATGCCGAGGTT	85
Qy	103	CCTGGAAATCCCTGGGGTCTTCGATCCGGAAATCTTACCCGGCATCCCTCCTAGGAGGG	162
Db	86	CCTGGAAATCCCTGGGGTCTTCGATCCGGAAATCTTACCCGGCATCCCTCCTAGGAGGG	145
Qy	163	ATTATATTATTTTTCCTTTTAACTGGAAGAGAGAAACAAGTTGTCTTTCCCC	222
Db	146	ATTATATTATTTTTCCTTTTAACTGGAAGAGAGAAACAAGTTGTCTTTCCCC	205
Qy	223	CTTCTCTGCTAAACCCATGATATACTGAATTAAGCGGCTCAGGGCTTCCCGCGT	282
Db	206	CTTCTCTGCTAAACCCATGATATACTGAATTAAGCGGCTCAGGGCTTCCCGCGT	265
Qy	283	GGAGCTCGAGGCCACATCTGCCTGATTCGCCGAGCCCGGAGGTTAGTCGAG	342
Db	266	GGAGCTCGAGGCCACCATCTGCCTGATTCGCCGAGCCCGGAGGTTAGTCGAG	325
Qy	343	TCTCTCGCGCGGGAGGATGGTGGCCGAGCCGGGAGCCCGGGCCCGCGGAGC	402
Db	326	TCTCTCGCGCGGGAGGATGGTGGCCGAGCCGGGAGCCCGGGCCCGCGGAGC	385
Qy	403	CGGCTCGGTGCCACCCAGCCGGGGTAGATGCTGCCCTCGCCAGCGCTGAGTGAC	462
Db	386	CGGCTCGGTGCCACCCAGCCGGGGTAGATGCTGCCCTCGCCAGCGCTGAGTGAC	445
Qy	463	ACCATGGAGACCTTCTGGTGGCTGCTAGCGTTTGGCATGGGCTTCCCGTGGTGCAC	522
Db	446	ACCATGGAGACCTTCTGGTGGCTGCTAGCGTTTGGCATGGGCTTCCCGTGGTGCAC	505
Qy	523	GGCTGCCCAAGTACTGTGTGTCGCCAGAACTGTGTGAGTCACTGGGAGCCCTGTGCC	582
Db	506	GGCTGCCCAAGTACTGTGTGTCGCCAGAACTGTGTGAGTCACTGGGAGCCCTGTGCC	565
Qy	583	TCCAGGGGCTGCTCTTTGTACCCCTGATATGACCCGCGGACATGGAGCTGGCCCTG	642
Db	566	TCCAGGGGCTGCTCTTTGTACCCCTGATATGACCCGCGGACATGGAGCTGGCCCTG	625
Qy	643	GGCGCACTTCATATCCACATCAGCCGCCAGACTTGGCAACATCAGCGGGCTGGT	702
Db	626	GGCGCACTTCATATCCACATCAGCCGCCAGACTTGGCAACATCAGCGGGCTGGT	685
Qy	703	GACCTGACCTGTCCAGGAACACCATFCAGCCACATCAGCCCTTTTCTTTCTGGACCTC	762
Db	686	GACCTGACCTGTCCAGGAACACCATFCAGCCACATCAGCCCTTTTCTTTCTGGACCTC	745
Qy	763	GAGAGCTCCGCTCCGCTGATCTTGAAGCAATCCGCTCCCAAGCTTTGGGAGGACACC	822
Db	746	GAGAGCTCCGCTCCGCTGATCTTGAAGCAATCCGCTCCCAAGCTTTGGGAGGACACC	805
Qy	823	CTCCGGGGCTGCTCAACCTGCAGCACCTTATCGTGAAACAAACAGCTGGCGGCATC	882

Db	806	CTCCGGGGCTGGTCAACCTGCAGCACCTTATCGTGAACAACAACAGCTGGCGGCATC	865
Qy	883	GCAGATGAGGCTTTGAGGACTTCTGTGACATGGAGGATCGAGCTCTCTTACAAAC	942
Db	866	GCAGATGAGGCTTTGAGGACTTCTGTGACATGGAGGATCGAGCTCTCTTACAAAC	925
Qy	943	AACCTCAATGGCTGCGGCTCCGCTGACGCGATCGTCAACCTCCACAGCTGAGC	1002
Db	926	AACCTCAATGGCTGCGGCTCCGCTGACGCGATCGTCAACCTCCACAGCTGAGC	985
Qy	1003	CTGGACCAACAACCTGTGGATCAATFCGCGGAGGACCTTTTGACAGCTGCAGAACTG	1062
Db	986	CTGGACCAACAACCTGTGGATCAATFCGCGGAGGACCTTTTGACAGCTGCAGAACTG	1045
Qy	1063	GCCCGCTGGATCAACCTCAATCGCTGCAGAACTGCGCCCTCATCCCACTTTTGCC	1122
Db	1046	GCCCGCTGGATCAACCTCAATCGCTGCAGAACTGCGCCCTCATCCCACTTTTGCC	1105
Qy	1123	CGCTCCAGGCTTCGGCTTTGACAGCCACACCTTTGCCCCACCCTTGCTTTAGTTTT	1182
Db	1106	CGCTCCAGGCTTCGGCTTTGACAGCCACACCTTTGCCCCACCCTTGCTTTAGTTTT	1165
Qy	1183	GGGGTAAACCCTTCACTGCAATTTGAGCTTCTCTGGCTGGAGGCTCGAGCGGAC	1242
Db	1166	GGGGTAAACCCTTCACTGCAATTTGAGCTTCTCTGGCTGGAGGCTCGAGCGGAC	1225
Qy	1243	GATGACTGGAAACCTGTGGCTCCAGGGGCTCAAGGGTGCCTACTTCTGGCATGTG	1302
Db	1226	GATGACTGGAAACCTGTGGCTCCAGGGGCTCAAGGGTGCCTACTTCTGGCATGTG	1285
Qy	1303	CGTAGGAGGAGTTGTGTGGAGCCGCTCTCATACCCAGCACACACAACTTTGGTG	1362
Db	1286	CGTAGGAGGAGTTGTGTGGAGCCGCTCTCATACCCAGCACACACAACTTTGGTG	1345
Qy	1363	GTTCGAGGGCCAGGGCCACACTCAAGTGAAAGCCATTGGGGACCCACAGCCCTT	1422
Db	1346	GTTCGAGGGCCAGGGCCACACTCAAGTGAAAGCCATTGGGGACCCACAGCCCTT	1405
Qy	1423	ATCCACTGGTAGCCCTCCGATGACCGCTGGTAGGAACTCTCAAGGACCGCTGTAT	1482
Db	1406	ATCCACTGGTAGCCCTCCGATGACCGCTGGTAGGAACTCTCAAGGACCGCTGTAT	1465
Qy	1483	GACAAATGGACCCCTGGACATTTTATCACACATCTAGGACAGTGTGCTTCACTGC	1542
Db	1466	GACAAATGGACCCCTGGACATTTTATCACACATCTAGGACAGTGTGCTTCACTGC	1525
Qy	1543	ATTGCTGCAATGCTCCGGAGAGCCACGCGCATGGTAGGTCTCCATCGTCCAGCTG	1602
Db	1526	ATTGCTGCAATGCTCCGGAGAGCCACGCGCATGGTAGGTCTCCATCGTCCAGCTG	1585
Qy	1603	CCACACTCAGCAACAGCACCAGCGCTGCAACCCCGCAAGTCCGCTCTCAGACATC	1662
Db	1586	CCACACTCAGCAACAGCACCAGCGCTGCAACCCCGCAAGTCCGCTCTCAGACATC	1645
Qy	1663	ACTGGCTCAGCAACAGCAGCGGGGAGTGGAGGAGTGGGGGGGAGAGCTCCAAA	1722
Db	1646	ACTGGCTCAGCAACAGCAGCGGGGAGTGGAGGAGTGGGGGGGAGAGCTCCAAA	1705
Qy	1723	AGCCCCCGGAAAGGGCTGTGTGTGTGCTGAAGTGACCACTCGGCCCTGTGCAAG	1782
Db	1706	AGCCCCCGGAAAGGGCTGTGTGTGTGCTGAAGTGACCACTCGGCCCTGTGCAAG	1765
Qy	1783	TGTFGTGTGACAAAGTCCAGCCCGGGTGAAGATGTACAGCTGCACTGCACTCTCT	1842
Db	1766	TGTFGTGTGACAAAGTCCAGCCCGGGTGAAGATGTACAGCTGCACTGCACTCTCT	1825
Qy	1843	GAGATGAGTACTGTATTTACAGGATGATCCAGCTTCCAAAGGCTTGTGTTCAAC	1902
Db	1826	GAGATGAGTACTGTATTTACAGGATGATCCAGCTTCCAAAGGCTTGTGTTCAAC	1885
Qy	1903	AACCTGTTGTCAGGACTGCTTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1962

Db 1886 AACCTGGTGTCTAGGACTGGTACTAGACTTGTGTGTCTGGCCATGTGGATGACACAGCC 1945
 Qy 1963 ACGACACTACGCGCCACCAACATCTGTGGCTGGCCAGTCTTTCACCAAGGCTGACTAC 2022
 Db 1946 ACGACACTACGCGCCACCAACATCTGTGGCTGGCCAGTCTTTCACCAAGGCTGACTAC 2005
 Qy 2023 CGCAGTGCAGTTCATGACAGCCAGATCTGTGGGGGACCAATGATCTGTGTCATCGGG 2082
 Db 2006 CGCAGTGCAGTTCATGACAGCCAGATCTGTGGGGGACCAATGATCTGTGTCATCGGG 2065
 Qy 2083 GCATCATCTGGCCACGCTGTGTCTTCATCGTCTCATCTCATCTGTGGGGGACCAATG 2142
 Db 2066 GCATCATCTGGCCACGCTGTGTCTTCATCGTCTCATCTCATCTGTGGGGGACCAATG 2125
 Qy 2143 TGCAACACAGGCCCCCAGCAAGATGGAGGGCCGTGAGCAATGTACTCGAGACC 2202
 Db 2126 TGCAACACAGGCCCCCAGCAAGATGGAGGGCCGTGAGCAATGTACTCGAGACC 2185
 Qy 2203 AACGGGCCAGCCAGCCCTCAAGCAGGACACAGCCGGGGCCCGCCGAGGGCCCG 2262
 Db 2186 AACGGGCCAGCCAGCCCTCAAGCAGGACACAGCCGGGGCCCGCCGAGGGCCCG 2245
 Qy 2263 CGAGAGTGGTGGCCAGCAGCTCTGGACTTCACCGCCAGCCTGGCCCGCCAGT 2322
 Db 2246 CGAAGTGGTGGCCAGCAGCTCTGGACTTCACCGCCAGCCTGGCCCGCCAGT 2305
 Qy 2323 GACTCTCTTCTCCAGCTCCCTGGGAGTGGGGAGTGGGGAGTGGGGAGTGGGGAGTGG 2382
 Db 2306 GACTCTCTTCTCCAGCTCCCTGGGAGTGGGGAGTGGGGAGTGGGGAGTGGGGAGTGG 2365
 Qy 2383 TGGAGATCCACCTCCCGCCCGCCCAAGCCAGCCCTTGACCCCTGTATGGGGCC 2442
 Db 2366 TGGAGATCCACCTCCCGCCCGCCCAAGCCAGCCCTTGACCCCTGTATGGGGCC 2425
 Qy 2443 TTCCTCTCCCTGAGCTCAAGAGTCAAGAAAGGAGGAGTCTGGACTCCAGACTCCA 2502
 Db 2426 TTCCTCTCCCTGAGCTCAAGAGTCAAGAAAGGAGGAGTCTGGACTCCAGACTCCA 2485
 Qy 2503 GCCGGAGAGGGGCTGGAGCTCGGCCCGGGCCACCTCGGACCGAGGACTGCTG 2562
 Db 2486 GCCGGAGAGGGGCTGGAGCTCGGCCCGGGCCACCTCGGACCGAGGACTGCTG 2545
 Qy 2563 GGGCCCTTGGCCCGCCAGGAGCTGCTCCCTTCCCTTGGAGGGCAAGGCCAAA 2622
 Db 2546 GGGCCCTTGGCCCGCCAGGAGCTGCTCCCTTCCCTTGGAGGGCAAGGCCAAA 2605
 Qy 2623 CGCAGCCTCTTGCACATGGGACTTGTGTGTGGCCGGGGGGTCTGTC 2682
 Db 2606 CGCAGCCTCTTGCACATGGGACTTGTGTGTGGCCGGGGGGTCTGTC 2665
 Qy 2683 GGGGCTACAGTCTCTCGAAGTCTCGAACATCTGGACGAGGAGGAGGAGGAGGAGG 2742
 Db 2666 GGGGCTACAGTCTCTCGAAGTCTCGAACATCTGGACGAGGAGGAGGAGGAGGAGG 2725
 Qy 2743 AACGGCATCTTGGCCCTTGGAGAGTGAACCTGGTGGGGCCCGGGGACTTTTGGC 2802
 Db 2726 AACGGCATCTTGGCCCTTGGAGAGTGAACCTGGTGGGGCCCGGGGACTTTTGGC 2785
 Qy 2803 AGCTCCGAATGGTGTAGAGACAGCCGCTAG 2835
 Db 2786 AGCTCCGAATGGTGTAGAGACAGCCGCTAG 2818

RESULT 4
 AAI58215
 ID AAI58215 standard; cDNA; 2818 BP.
 XX
 AC AAI58215;
 XX
 DT 22-OCT-2001 (first entry)
 XX
 DE Human polynucleotide SEQ ID NO 418.
 XX

Human; neutropic; immunosuppressant; cytostatic; gene therapy; cancer; peripheral nervous system; neuropathy; central nervous system; CNS; Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic; amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic; chemokinetic; thrombolytic; drug screening; arthritis; inflammation; leukaemia; ss.
 Homo sapiens.
 WO200153312-A1.
 26-JUL-2001.
 26-DEC-2000; 2000WO-US34263.
 21-JAN-2000; 2000US-0488725.
 25-APR-2000; 2000US-0552317.
 09-JUL-2000; 2000US-0598042.
 19-JUL-2000; 2000US-0620312.
 03-AUG-2000; 2000US-0653450.
 14-SEP-2000; 2000US-0662191.
 19-OCT-2000; 2000US-0693036.
 29-NOV-2000; 2000US-0727344.
 (HYSE-) HYSEQ INC.
 Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
 Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
 Zhao QA, Zhou P, Goodrich R, Demanac RT;
 WPI: 2001-442253/47.
 P-PSDB; AAM39059.
 Novel nucleic acids and polypeptides, useful for treating disorders such as central nervous system injuries -
 Claim 1; SEQ ID NO 418; 10078pp; English.
 The invention relates to human nucleic acids (AAI57798-AAI61369) and the encoded polypeptides (AAM38642-AAM42213) with neutropic immunosuppressant and cytostatic activity. The polynucleotides are useful in gene therapy. A composition containing a polypeptide or polynucleotide of the invention may be used to treat diseases of the peripheral nervous system, such as peripheral nervous injuries, peripheral neuropathy and Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis, and Shy-Drager Syndrome. Other uses include the utilisation of the activities such as: Immune system suppression, Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic and thrombolytic activity, cancer diagnosis and therapy, drug screening, assays for receptor activity, arthritis and inflammation, leukaemias and C.N.S disorders.
 Note: The sequence data for this patent did not form part of the printed specification.
 Sequence 2818 BP; 523 A; 933 C; 825 G; 537 T; 0 other;
 Query Match 88.4%; Score 2780.2; DB 22; Length 2818;
 Best Local Similarity 99.7%; Pred. No. 0;
 Matches 2785; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
 Qy 43 CACAACCCGGCAGGCTCGTGAACCTTGAAGACACCCACATTCACAAGATGCCCGAGGTT 102
 Db 26 CCCCCCCCCCTGGCTCGTGAACCTTGAAGACACCCACATTCACAAGATGCCCGAGGTT 85'
 Qy 103 CTTGGAAATGCTGGGGTTCTTCATCCGAAAATCTTACCGGATCTCTTAGGGAGG 162
 Db 86 CTTGGAAATGCTGGGGTTCTTCATCCGAAAATCTTACCGGATCTCTTAGGGAGG 145
 Qy 163 ATTAT 222
 Db 146 ATTAT 205

```

QY 223 CTTCTTCTTAACGCCATCGATATAACTGAAATAAGGGCTCAGGCTTCCCGCGT 282
Db 206 CTTCTTCTTAACGCCATCGATATAACTGAAATAAGGGCTCAGGCTTCCCGCGT 265
QY 283 GGAGCTCCGAGGCCACCATCTCCCTGCAATCCCGGAGCCCGGAGGGTTAGTCGAG 342
Db 266 GGAGCTCCGAGGCCACCATCTCCCTGCAATCCCGGAGCCCGGAGGGTTAGTCGAG 325
QY 343 TCTGTCGCGGGGGAAGGATGCGFTGCCAGCCGGGAGCCCGGGGCCCCCGGGAGC 402
Db 326 TCTGTCGCGGGGGAAGGATGCGFTGCCAGCCGGGAGCCCGGGGCCCCCGGGAGC 385
QY 403 CGGCTCGTGCACCCAGCCGGGGTAGCTGCTCGCCAGCCAGCCGCTGAGTCACCAG 462
Db 386 CGGCTCGTGCACCCAGCCGGGGTAGCTGCTCGCCAGCCGGGAGCCCGGGGCCCCCGGGAGC 445
QY 463 ACCATGGAGACCCTGCTGGTGGCCTGCTAGGCTTTGGCATGGCCGTTTCCCGTGGCAG 522
Db 446 ACCATGGAGACCCTGCTGGTGGCCTGCTAGGCTTTGGCATGGCCGTTTCCCGTGGCAG 505
QY 523 GCCTGCCCAAGFPACTGTGTGCCAGAAATCTGTCTGAGTCACTGGGACCCCTGTGCCCC 582
Db 506 GCCTGCCCAAGFPACTGTGTGCCAGAAATCTGTCTGAGTCACTGGGACCCCTGTGCCCC 565
QY 583 TCCAGGGCTGCTTTTGTACCCTGATATTTGACCCGGGACAGTGGAGCTGGCCCTG 642
Db 566 TCCAGGGCTGCTTTTGTACCCTGATATTTGACCCGGGACAGTGGAGCTGGCCCTG 625
QY 643 GCGGCAACTTTCATCATCCACATCAGCCCGGAGGCTTTGCCAAATGACGGGGCTGGTG 702
Db 626 GCGGCAACTTTCATCATCCACATCAGCCCGGAGGCTTTGCCAAATGACGGGGCTGGTG 685
QY 703 GACTGACCTGTCCAGGAAACCATCAGCCATCCAGCCCTTTCCTTTCGACCTC 762
Db 686 GACTGACCTGTCCAGGAAACCATCAGCCATCCAGCCCTTTCCTTTCGACCTC 745
QY 763 GAGGCTCCGCTCCCTGCACTTTGAGCAATCGGCTGCAAGCCTTGGGAGGACACC 822
Db 746 GAGGCTCCGCTCCCTGCACTTTGAGCAATCGGCTGCAAGCCTTGGGAGGACACC 805
QY 823 CTCGGGGCTGTGTAACCTGACGACCTTTCGTGAACAAACACAGCTGGGCGGATC 882
Db 806 CTCGGGGCTGTGTAACCTGACGACCTTTCGTGAACAAACACAGCTGGGCGGATC 865
QY 883 GCAGTGGCTTTGAGGACTTCTGCTGACATTTGAGGATCTGGACCTTCTCTACAAC 942
Db 866 GCAGTGGCTTTGAGGACTTCTGCTGACATTTGAGGATCTGGACCTTCTCTACAAC 925
QY 943 AACCTCATGGCCCTGGGCTGGGACTCGGTGGAGCAATGGTCAACCTCCACCCAGCTGAGC 1002
Db 926 AACCTCATGGCCCTGGGCTGGGACTCGGTGGAGCAATGGTCAACCTCCACCCAGCTGAGC 985
QY 1003 CTGGACCACAACTGCTGGATACATCGCCGAGGCACTTTGAGACACTGCAAGAACTG 1062
Db 986 CTGGACCACAACTGCTGGATACATCGCCGAGGCACTTTGAGACACTGCAAGAACTG 1045
QY 1063 GCGCGCTGATCTCACTCCMATCCGCTGACAGAGCTCCCGTATCCCATCTTTGCC 1122
Db 1046 GCGCGCTGATCTCACTCCMATCCGCTGACAGAGCTCCCGTATCCCATCTTTGCC 1105
QY 1123 CGCTCCAGGCTTGGGCTTTCAGACCCACACCTTTCGCCCCACCTTTGTCTTTAGTTT 1182
Db 1106 CGCTCCAGGCTTGGGCTTTCAGACCCACACCTTTCGCCCCACCTTTGTCTTTAGTTT 1165
QY 1183 GGGGTAAACCACTTACTGCAATTTGAGTCTCTGGCTGCGGAGCTCGAGGGGAC 1242
Db 1166 GGGGTAAACCACTTACTGCAATTTGAGTCTCTGGCTGCGGAGGCTCGAGGGGAC 1225
QY 1243 GATGACCTGAAACTGCTGGCTCCCGAGGGCTCAAGGCTGCTACTTCTGCAATGTG 1302
Db 1226 GATGACCTGAAACTGCTGGCTCCCGAGGGCTCAAGGCTGCTACTTCTGCAATGTG 1285
QY 1303 CGTGAGGAGGAGTTTGTGGGAGCCGCTCTCATCACCCAGCACACAAAGTTGCTG 1362

```

```

Db 1286 CGTGAGGAGAGTTTGTGGAGCCGCTCTCATCCCCAGCACACACAAGTTGCTG 1345
QY 1363 GTTCTGAGGGCCAGGGCCACACTCAAGTGAAGCCATTTGGGACCCAGCCCTT 1422
Db 1346 GTTCTGAGGGCCAGGGCCACACTCAAGTGAAGCCATTTGGGACCCAGCCCTT 1405
QY 1423 ATCCACTGGGTAGCCCCCGATGACCCGCTGTTAGGAACTCCTCAAGACCCGTTCTAT 1482
Db 1406 ATCCACTGGGTAGCCCCCGATGACCCGCTGTTAGGAACTCCTCAAGACCCGTTCTAT 1465
QY 1483 GACAAATGGCCACCTTGAATCTTTCATCACACATCTCAGGACAGTGTGTGTTCACCTGC 1542
Db 1466 GACAAATGGCCACCTTGAATCTTTCATCACACATCTCAGGACAGTGTGTGTTCACCTGC 1525
QY 1543 ATTCTGCCAATGTCGCGGAGGACCGCCATGTCGAGGCTTCATCTGTCAGGCTG 1602
Db 1526 ATTCTGCCAATGTCGCGGAGGACCGCCATGTCGAGGCTTCATCTGTCAGGCTG 1585
QY 1603 CCACACCTCAGCAACAGCACCGCCTGACCCCCCAAGTCCCGCTCTCAGACATC 1662
Db 1586 CCACACCTCAGCAACAGCACCGCCTGACCCCCCAAGTCCCGCTCTCAGACATC 1645
QY 1663 ACTGGCTCCAGCAACAGCACCGCCTGACCGCCAGTGGGGGGGAGAGCTCCCAAA 1722
Db 1646 ACTGGCTCCAGCAACAGCACCGCCTGACCGCCAGTGGGGGGGAGAGCTCCCAAA 1705
QY 1723 AGCCCCCGAAGCGGCTGTGTTGCTGAAGTGACCCACCTCGGCTCTGTCAG 1782
Db 1706 AGCCCCCGAAGCGGCTGTGTTGCTGAAGTGACCCACCTCGGCTCTGTCAG 1765
QY 1783 TGGTCTGTCAGCAAGTCAAGCAGCCCGGTGAAGTGTACAGCTGCAGTCAAACTGCTCT 1842
Db 1766 TGGTCTGTCAGCAAGTCAAGCAGCCCGGTGAAGTGTACAGCTGCAGTCAAACTGCTCT 1825
QY 1843 GACGATGACTGATTTACAGSAGTCCAGCCTCACAAGGCTTCGTGTCAAC 1902
Db 1826 GACGATGACTGATTTACAGSAGTCCAGCCTCACAAGGCTTCGTGTCAAC 1885
QY 1903 AACCTGGTTCAGGACTGGCTGACCTGTTGTGTGTGGCCATGTGGGATGACACAGCC 1962
Db 1886 AACCTGGTTCAGGACTGGCTGACCTGTTGTGTGTGGCCATGTGGGATGACACAGCC 1945
QY 1963 ACGACACTCAGCCCAACAACTGTGGCTGCGCCAGTTCACCAAGGCTGACTAC 2022
Db 1946 ACGACACTCAGCCCAACAACTGTGGCTGCGCCAGTTCACCAAGGCTGACTAC 2005
QY 2023 CCGGATGCCAGTCCATCAGACAGCAGATTTCTGGCCGACCATGATCTGTCTATCGGG 2082
Db 2006 CCGGATGCCAGTCCATCAGACAGCAGATTTCTGGCCGACCATGATCTGTCTATCGGG 2065
QY 2083 GGCATCATCGTGGCCACGCTGCTGTTTCATCTGTCATCTCATGGTGGCTTCAAGGTC 2142
Db 2066 GGCATCATCGTGGCCACGCTGCTGTTTCATCTGTCATCTCATGGTGGCTTCAAGGTC 2125
QY 2143 TGCAACAGGAGGCCCCAGCAAGATGGAGCCGCGTGAAGAAATGTGTACTGCGAGCC 2202
Db 2126 TGCAACAGGAGGCCCCAGCAAGATGGAGCCGCGTGAAGAAATGTGTACTGCGAGCC 2185
QY 2203 AAGGGCCAGCCACCGCTCCAAAGCAGCAGCAGCCAGCCGCGGCCCCCGGAGGCCG 2262
Db 2186 AAGGGCCAGCCACCGCTCCAAAGCAGCAGCAGCAGCCAGCCGCGGCCCCCGGAGGCCG 2245
QY 2263 CCGAAGGTGGTGGCGCAACAGCTCTGGAATTCACCGCCAGCTGGCCCGCGGCGAGT 2322
Db 2246 CCGAAGGTGGTGGCGCAACAGCTCTGGAATTCACCGCCAGCTGGCCCGCGGCGAGT 2305
QY 2323 GACTCTCTTCTTCAGCTCCCTGGCAGTGGGGCTCGGGGCTGAGAGGGGCCCC 2382
Db 2306 GACTCTCTTCTTCAGCTCCCTGGCAGTGGGGCTCGGGGCTGAGAGGGGCCCC 2365
QY 2383 TGGAGGATCCACCTTCCCGCCCGCCAGCCGCTTACCCGCTTGTGGGGG 2442

```

Db 2366 TGAGATCCACCCCTCCCGCCCGCCCAAGCCAGCCCTTGACCGCCCTGATGGGGCC 2425
 Qy 2443 TTCGCTCCCTGACCTCAAGAGTCAAGAAAGGAGGAGTCTGGACTCCAGGACTCCA 2502
 Db 2426 TTCGCTCCCTGACCTCAAGAGTCAAGAAAGGAGGAGTCTGGACTCCAGGACTCCA 2485
 Qy 2503 GCCGGGAGAGGGCTGGACGTCGGCCCGGGGCCACCCTCGGACCGAGACCTGCTG 2562
 Db 2486 GCCGGGAGAGGGCTGGACGTCGGCCCGGGGCCACCCTCGGACCGAGACCTGCTG 2545
 Qy 2563 GGGCCCCCTCGCCCGCCGACGAGCTCCCTCCCTGCTGGAGGCAAGCCAAA 2622
 Db 2546 GGGCCCCCTCGCCCGCCGACGAGCTCCCTCCCTGCTGGAGGCAAGCCAAA 2605
 Qy 2623 CGCAGCCTCCTCGACATGGGGACTTTGCTGCGCGCGGGAGGGTCTGTCGG 2682
 Db 2606 CGCAGCCTCCTCGACATGGGGACTTTGCTGCGCGCGGGAGGGTCTGTCGG 2665
 Qy 2683 GCGCGTACAGTCTCCTCGAAGGTCTCGAACATCTGGACGAAGCCCTCTCTGTC 2742
 Db 2666 GCGCGTACAGTCTCCTCGAAGGTCTCGAACATCTGGACGAAGCCCTCTCTGTC 2725
 Qy 2743 AACGCATCTCTTGCCTTTTCAGGAGAGTACCTGGTGGGGCCCGGGGACTTTTGGC 2802
 Db 2726 AACGGATCTCTTGCCTTTTCAGGAGAGTACCTGGTGGGGCCCGGGGACTTTTGGC 2785
 Qy 2803 AGCTCCGAATGGGTGATGAGAGCAGCGGTCTAG 2835
 Db 2786 AGCTCCGAATGGGTGATGAGAGCAGCGGTCTAG 2818

RESULT 5

AAS77723
 ID AAS77723 standard; cDNA; 969 BP.
 XX AAS77723;
 AC AAS77723;
 XX
 DT 13-FEB-2002 (first entry)
 XX
 DE DNA encoding novel human diagnostic protein #13527.
 XX
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200175067-A2.
 XX
 PD 11-OCT-2001.
 XX
 XX 30-MAR-2001; 2001WO-US08631.
 XX
 PR 31-MAR-2000; 2000US-0540217.
 PR 23-AUG-2000; 2000US-0649167.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 XX Drmanac RT, Liu C, Tang YT;
 XX
 DR WPI; 2001-639362/73.
 DR P-PSDB; ABG13536.
 XX
 XX New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity -
 XX
 PS Claim 1; SEQ ID No 13527; 103pp; English.
 XX
 CC The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The

CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations in
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. AAS64197-AAS94564 represent novel human
 CC diagnostic coding sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX

SQ Sequence 969 BP; 166 A; 330 C; 316 G; 157 T; 0 other;
 Query Match 30.8%; Score 969; DB 23; Length 969;
 Best Local Similarity 100.0%; Pred. No. 1.4e-194;
 Matches 969; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1867 ATGATCCAGGCTCCAAACAAGGCTTCGTGGTCAACAACCTGGTTCAGGGACTGGCTAC 1926
 Db 1 ATGATCCAGGCTCCAAACAAGGCTTCGTGGTCAACAACCTGGTTCAGGGACTGGCTAC 60
 Qy 1927 GACTTGTGTGTCGGCCATGCGGATGACACAGCCAGCAGACTCACCAGCCACCACATC 1986
 Db 61 GACTTGTGTGTCGGCCATGCGGATGACACAGCCAGCAGACTCACCAGCCACCACATC 120
 Qy 1987 GTGGCTGGCCCGCTTCTTCAACAAGGCTGACTACCCCGCAGTGCAGTCCATGCACAGC 2046
 Db 121 GTGGCTGGCCCGCTTCTTCAACAAGGCTGACTACCCCGCAGTGCAGTCCATGCACAGC 180
 Qy 2047 CAGATTCTGGGCGCACCATGATCTGTGTAFCGGGGGCACTCATCTGGCCACGTCGTC 2106
 Db 181 CAGATTCTGGGCGCACCATGATCTGTGTAFCGGGGGCACTCATCTGGCCACGTCGTC 240
 Qy 2107 GTCTTCACTGCTCATCTGATGGTGTACAGGTCTCAAGGTCTCAACACAGAGCCCGCCAGCAG 2166
 Db 241 GTCTTCACTGCTCATCTGATGGTGTGTACAGGTCTGTCAACACAGAGCCCGCCAGCAG 300
 Qy 2167 ATGGCAGGCGGCGTGGCAATGTGTACTCGCAGACCAACGCGGCCCGCCAGCCCTCCA 2226
 Db 301 ATGGCAGGCGGCGTGGCAATGTGTACTCGCAGACCAACGCGGCCCGCCAGCCCTCCA 360
 Qy 2227 AGCAGCGCACACAGCCGGGCGCCCGCCAGGGCCCGCCGAAAGGTGGTGGCAACGAG 2286
 Db 361 AGCAGCGCACACAGCCGGGCGCCCGCCAGGGCCCGCCGAAAGGTGGTGGCAACGAG 420
 Qy 2287 CTCTTGGACTTACCCGCGGAGTGGCGCGGCGCGGCGCGGCGCGGCGCGGCGCGGCGCGG 2346
 Db 421 CTCTTGGACTTACCCGCGGAGTGGCGCGGCGCGGCGCGGCGCGGCGCGGCGCGGCGCGG 480
 Qy 2347 GGCAGTGGGGAGGCTGCGGGGTGGAGCGGGCCCGCTGGAGGATCCACCCCTCCGCGCCCG 2406
 Db 481 GGCAGTGGGGAGGCTGCGGGGTGGAGCGGGCCCGCTGGAGGATCCACCCCTCCGCGCCCG 540
 Qy 2407 CGCCCCAAGCCAGCCTTGACCGCCTGATGGGGCGCTTCGCTCCCTGGACCTCAAGAGT 2466
 Db 541 CGCCCCAAGCCAGCCTTGACCGCCTGATGGGGCGCTTCGCTCCCTGGACCTCAAGAGT 600
 Qy 2467 CAGAGAAAGGAGGAGCTCTGGACTCCAGGACTCCAGCGGGGAGGGGCTGGGCGTCG 2526
 Db 601 CAGAGAAAGGAGGAGCTCTGGACTCCAGGACTCCAGCGGGGAGGGGCTGGGCGTCG 660
 Qy 2527 GCCCGGGCCACCACTCGGACCGAGGACCTGTGGGGCCCGCTGGCGCCCGCCCGCCAGG 2586
 Db 661 GCCCGGGCCACCACTCGGACCGAGGACCTGTGGGGCCCGCTGGCGCCCGCCCGCCAGG 720
 Qy 2587 AGCCTGCTCCCTTCCGTTGGAGGGCAAGGGCCAAACGACCACTCCTTCGACATGGGG 2646

Db 721 ACCCTGCTCCCTTCGCTGGAGGCAAGCCAAACGACCCACTCCTTCACATGGG 780
 Qy 2647 GACTTTGCTGTCGGCGGCGGAGGGTCTGTCGGCGGCTACAGTCCCTCCGGAAG 2706
 Db 781 GACTTTGCTGTCGGCGGCGGAGGGTCTGTCGGCGGCTACAGTCCCTCCGGAAG 840
 Qy 2707 GTCCTGAACATCGACGAGCGGACCTCTCTGTCACAGGCACTCTTGCCTTTGAG 2766
 Db 841 GTCCTGAACATCGACGAGCGGACCTCTCTGTCACAGGCACTCTTGCCTTTGAG 900
 Qy 2767 GAGAGTGACCTGGTGGGGCCCGGGGACTTTTGGCAGCTCCGAATGGTGTATGAGAGC 2826
 Db 901 GAGAGTGACCTGGTGGGGCCCGGGGACTTTTGGCAGCTCCGAATGGTGTATGAGAGC 960
 Qy 2827 ACGGTCTAG 2835
 Db 961 ACGGTCTAG 969

RESULT 6
 AAI60001
 ID AAI60001 standard; cDNA; 771 BP.
 XX
 AC AAI60001;
 XX
 DT 22-OCT-2001 (first entry)
 XX
 DE Human polynucleotide SEQ ID NO 3990.
 XX
 KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
 KW peripheral nervous system; neuropathy; central nervous system; CNS;
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
 KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
 KW leukaemia; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200153312-A1.
 XX
 PD 26-JUL-2001.
 XX
 PF 26-DEC-2000; 2000MO-US4263.
 XX
 PR 21-JAN-2000; 2000US-0488725.
 PR 25-APR-2000; 2000US-0552317.
 PR 09-JUL-2000; 2000US-0598042.
 PR 19-JUL-2000; 2000US-0620312.
 PR 03-AUG-2000; 2000US-0653450.
 PR 14-SEP-2000; 2000US-0662191.
 PR 19-OCT-2000; 2000US-0693036.
 PR 29-NOV-2000; 2000US-0727344.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
 PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
 PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
 XX
 DR WPI: 2001-442253/47.
 DR P-PSDB; AAM40845.
 XX
 PT Novel nucleic acids and polypeptides, useful for treating disorders
 PT such as central nervous system injuries -
 XX
 PS Claim 1; SEQ ID NO 3990; 10078pp; English.
 XX
 CC The invention relates to human nucleic acids (AAI57798-AAI61369) and
 CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,
 CC immunosuppressant and cytostatic activity. The polynucleotides are useful
 CC in gene therapy. A composition containing a polypeptide or polynucleotide
 CC of the invention may be used to treat diseases of the peripheral nervous

CC system, such as peripheral nervous injuries, peripheral neuropathy and
 CC localised neuropathies and central nervous system diseases, such as
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 CC utilisation of the activities such as; Immune system suppression,
 CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening and
 CC assays for receptor activity, arthritis and inflammation, leukaemia and
 CC C.N.S disorders.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification.
 XX
 SQ Sequence 771 BP; 144 A; 259 C; 197 G; 171 T; 0 other;
 Query Match 21.6%; Score 678.8; DB 22; Length 771;
 Best Local Similarity 95.8%; Pred. No. 1.8e-133;
 Matches 730; Conservative 0; Mismatches 27; Indels 5; Gaps 3;

Qy 529 CCCAAGTACTGTCTGCCAGAAATGTCGTAGTCACTGGGGACCCCTGCCCCCTCAAG 588
 Db 8 CCCCAAGACTGTCTGCCAGAAATGTCGTAGTCACTGGGGACCCCTGCCCCCTCAAG 67
 Qy 589 GGGCTGCTCTTTACCCCTGATATTGACCGGGGACAGTGGAGCTGCCCTGGCGGC 648
 Db 68 GGGCTGCTCTTTACCCCTGATATTGACCGGGGACAGTGGAGCTGCCCTGGCGGC 127
 Qy 649 AACTTCATCATCCACATCAGCCGACAGCTTTGCCAAATGACGGGGCTGGTGGACTG 708
 Db 128 AACTTCATCATCCACATCAGCCGACAGCTTTGCCAAATGACGGGGCTGGTGGACTG 187
 Qy 709 ACCCTGTCAGGAAACACCATCAGCCATCCAGCCCTTTCCCTTTGGACTCGAGAGC 768
 Db 188 ACCCTGTCAGGAAACACCATCAGCCATCCAGCCCTTTCCCTTTGGACTCGAGAGC 247
 Qy 769 CTCGGCTCCCTGATCTTGACAGCAATCGGCTGCAAGCCCTTTGGGGAGCACCCCTCGG 828
 Db 248 CTCGGCTCCCTGATCTTGACAGCAATCGGCTGCAAGCCCTTTGGGGAGCACCCCTCGG 307
 Qy 829 GGCCTGGTCAACTGCGACACCTTATCGTGAACAACACCACTGGCGGCATCGCAGAT 888
 Db 308 GGCCTGGTCAACTGCGACACCTTATCGTGAACAACACCACTGGCGGCATCGCAGAT 367
 Qy 889 GAGGCTTTTGGAGCTTCTGTCGATTTGGAGGATCTGGACCTCTCTCAACAACCTC 948
 Db 368 GAGGCTTTTGGAGCTTCTGTCGATTTGGAGGATCTGGACCTCTCTCAACAACCTC 427
 Qy 949 CAT-GGCCTGCGCTGGACTCCGT-GCGACGCAATGG---TCAACCTCCACAGCTGAGCC 1003
 Db 428 CATGGGCTTCCGCTGGACTCCGTGGCGACGCAATGGGTTCAACCTCCACAGCTGAGCC 487
 Qy 1004 TGGACCAACCTGCTGGATCACATCGCGGAGGCAACCTTTGCAGACCTTGCAAGAACTGG 1063
 Db 488 TGGACCAACCTGCTGGATCACATCGCTCGAGGACACCTTTGCAGACCTTGCAAGAACTGG 547
 Qy 1064 CCGGCTGGATCTCACTCCAAATCGGCTGCAGAAAGTCCCTCATTCCTTTGGCC 1123
 Db 548 CCGGCTGGATCTCACTCCAAATCGGTTGCATAAAGTCCCTCATTCCTTTGGCC 607
 Qy 1124 GCTCCAGGCTTCCGCTTTGACAGCCACACCTTTGCCCCACCTTGTCTTTAGTTTGG 1183
 Db 608 GCTCCAGGCTTCCGCTTTGACAGACACACCTTTGCCCCACCTTGTCTTTATTTTGG 667
 Qy 1184 GGGTAAACCCACTTCACTGCAATTTGAGGCTTCTCTGGCTGGGAGGCTCGAGCGGGAGC 1243
 Db 668 GGGTAAACCCACTTCACTGCAATTTGAGGCTTCTCTGGCTGGGAGGCTCGATCGGGAGC 727
 Qy 1244 ATGACCTGGAAACCTGTGGTCCCGAGGGGGCTCAAGGTC 1285
 Db 728 AAACCTGGAAACCTTTGGTTCCTCCAGTGGGCGCTCAAGGGTC 769

RESULT 7
 AAH25568

AAH25568 standard; DNA; 1887 BP.	
AAH25568;	
05-SEP-2001 (first entry)	
Nucleotide sequence of an interferon omega-1 like protein NOV2.	
Interferon omega-1 like protein; Nov2; membrane bound protein; secreted protein; spermatogenesis; male infertility; neoplasia; blood circulation; immunological disorder; autoimmune disease; inflammatory disease; cardiovascular disease; metabolic disease; cancer; viral infection; acute lymphoblastic leukemia; glioma; neurological disease; neurodegenerative disorder; Alzheimer's disease; Parkinson's disorder; hematopoietic disorder; neurodegenerative disorder; immune disorder; hematopoietic disorder; ss.	
Homo sapiens.	
Key	Location/Qualifiers
CDS	1..1887
	/*tag= a
	/product= "interferon omega-1 like protein NOV2"
WO200142471-A2.	
14-JUN-2001.	
08-DEC-2000; 2000WO-US33463.	
09-DEC-1999; 99US-0169887.	
10-DEC-1999; 99US-0170230.	
07-DEC-2000; 2000US-0170230.	
(CURA-) CURAGEN CORP.	
Prayaga SK, Shinkets RA;	
WPI; 2001-381691/40.	
P-PSDB; AAB84469.	
Novel polypeptides designated as NOV polypeptides, useful in detection, prevention and treatment of a broad range of pathological states	
Claim 8; Page 15-16; 92pp; English.	
The present sequence encodes an interferon omega-1 like protein, designated NOV2. The protein is a membrane bound and secreted protein. NOV polypeptides and polynucleotides are useful for treating or preventing a NOV-associated disorder, and in the manufacture of a medicament for treating a syndrome associated with a human disease such as NOV-associated disorder. They are also useful for treating disorders or syndromes including those involved in development, differentiation and activation of thymic immune cells, pathologies related to spermatogenesis and male infertility, diagnosis of several human neoplasias, in disease or pathologies of cells in blood circulation such as red blood cells and platelets, various immunological disorders and/or pathologies, autoimmune and inflammatory diseases, cardiovascular diseases, metabolic diseases, cancer growth and metastasis, viral infections, cancer therapy, acute lymphoblastic leukemia, in gliomas, neurological diseases, neurodegenerative disorders, Alzheimer's disease, Parkinson's disorder, and hematopoietic disorders. NOV polypeptides and polynucleotides are also useful for treating neurodegenerative disorders, immune disorders and hematopoietic disorders.	
Sequence 1887 BP; 259 A; 736 C; 577 G; 315 T; 0 other;	
Query Match	20.2%; Score 634.6; DB 22; Length 1887;
Best Local Similarity	61.3%; Pred. No. 4.6e-124;
Matches 1079; Conservative	0; Mismatches 664; Indels 16; Gaps 3;
507 GTTGGCGGTGCGACCGCTGCCCAAGTACTGTGTCTCCAGAAATCTCTGTGACT 566	

63	GTGAGCCACACCCAGCCATGTCCCGCGCTGCCAGACACAGTTCGCTGCCCT	122
567	GGGACCCTGTGCTCCAAAGGGGCTGCTTTGTACCCCTGATATTGACCGGGAC	626
123	AAGCGTGTGTCGCCAGGGGAGGCTCCTGTTGTCGCCACCTCGCTGAGACCGCGG	182
627	AGTGGAGTGGCCCTGGCGGCAACTTCATATCCACATCCAGCCGCGAGACTTTGCCAA	686
183	AGCCGAGTCCGCTGGCAGACAACTTTCATCGCTCGGTCGCGCGGAGCTGGCCAA	242
687	CATGAGGGGCTGGTGGACCTGTCCAGAACACCATCAGCCACATCCAGCCCTT	746
243	CATGACGGCTGTGCATCTGAGCTGTCGGCGAACACCATCCGCGAGCTGGCTCGCG	302
747	TTCCCTTTCGACCTCGAGGCTCCGCTCCCTGATCTTACAGCAATCCGCTGCAAG	806
303	CGCCTTTCGCGACTGGCGGCTCGCTGCTGCACCTGGATGGCAACCGCTGACCTC	362
807	CCTGGGAGGACACCTCCGCGGCTGGTCAACCTGACGACCTTATCGTGAACAACA	866
363	ACTGGGAGGCGCAGCTGCGGGCTGGTCAACTTTGGCCACCTCATCTCAGCAACA	422
867	CGAGCTGGCGGCTCCAGATGAGGCTTTTGAGGACTTCTCTGTGACATTTGAGGATCT	926
423	CCAGCTGGCAGCGCTGGCGGCGGCTGGATGATTTGTGCCGAGACATCGAGGACCT	482
927	GGACCTCTCTACAAACCTCCATGGCTGCGGCTGGGACTCCGTGCGAGCATGGTCAA	986
483	CGACCTCTCTACAAACCTCGAGCTGCGGCTGGGAGGCTGGGGCGGCTGGGCAA	542
987	CCTCCACAGCTGAGCTGGACCAACCTGTGATCACATCGCGGAGGACCTTTGC	1046
543	CGTCAACAGTTGGGCTCGACCAACAACCTGTGCTGCTGCTGCGCGGCGCTTTTC	602
1047	AGACTGCAGAAACTGGCGGCTGGATCTCACTTCAACTCGGCTGCGAGGAGTGC	1106
603	CGCCCTGCACAAAGCTGGCGGCTGGATGACATGACCTCAACCGCCTGACCAACA	662
1107	TGATCCCATTTTGGCGCTCCCGCTTTGAGCGACACACCTTTGGCGCGCC	1166
663	CGACCCACTTCTCCCGCTGCCCCCTGCTGCCAGGCCCGGGGCTGCCCGCTCTGC	722
1167	CTTGTCTTTAGTTTTGGGGTAAACCCACTTCACTGCATTTGTGAGCTTCTCTGGG	1226
723	CCTGCTGCTGGCTTTGGCGGAAACCCCTGCACTGCAACTCGAGTGTGTGCTGCG	782
1227	GAGCTCGAGGGGAGATGACCTGGAAACTGTGGTCTCCCGAGGGGGCTCAAGGGT	1286
783	TGCGCTGGCGGGGAGGACGACTCGAGGCTTCGCGCTCCCGCCTGCTGGCGGCGG	842
1287	CTACTTCTGGCATGTGCTGAGGAGGTTGTTGTGCGAGCGGCTCTCATCACCAGCA	1346
843	CTACTTCTGGCGGGTGGCGAGGAGTTTGTCTGCGAGGCCCGCGTGTGACTCACCG	902
1347	CACACAAAGTTGCTGTTTGGAGGGCAGCGGCCACACTCAAGTGAAGACCAATTGG	1406
903	CTACCACTCTGGTGTGCGCGAGTTCGGCGGCTGCGCTGCGCTGCGCGGAGTGGG	962
1407	GGACCCAGCCCCCTTATCCACTGGGTAGCCCCGATACCCCTGTTAGGAACTCTTC	1466
963	GGACCCAGACCCCGTGTGGTGTGCGTGTGTCACCCAGGGCCGCTGTAGGCAACTCAAG	1022
1467	AAGGACCGCTGTATGACATGGACCCCTGGACATCTTATCACCACATCTCAGGACAG	1526
1023	CGTGCCTGGCGCTTCCCCTAATGGGACGCTGGAGTGTGTGGTACCAGCCGGGTGG	1082
1527	TGGTCCCTTCACTGCTTGGCAATGCTGCCGAGAGGGCCAGGGCCATGTTGAGGT	1586
1083	TGGCATCTTCACTGCTGCTTGGCCAAATGACAGCTGGGAGGCGCACAGCTGTGGAGCT	1142
1587	CTCCATGTGTCCAGCTGGCAGCCTCAGCAACAGCACAGCCAGCCGCA--CTGAC	1644

/product= "secreted polypeptide"

WO200166690-A2.
13-SEP-2001.
05-MAR-2001; 2001WO-US07143.
06-MAR-2000; 2000US-0187107.
13-MAR-2000; 2000US-0188916.
03-OCT-2000; 2000US-0236874.
03-OCT-2000; 2000US-0237846.
(SMIK) SMITHKLINE BEECHAM CORP.
(SMIK) SMITHKLINE BEECHAM PLC.
Agarwal P, Murdoch PR, Rizvi SK, Smith RF, Xiang Z;
WPI; 2001-570768/64.
P-PSDB; AAG67523.
Novel isolated secreted polypeptide useful for treating nervous and muscular diseases, gastrointestinal ulceration, coagulation and immune disorders, microbial diseases, inflammation and transplant rejection -
Claim 2; Page 49-50; 102pp; English.
The present sequence encodes a human secreted polypeptide. The secreted polypeptides and polynucleotides are useful for treating nervous and muscular diseases, for inhibiting tumour formation and metastasis, for treating gastrointestinal ulceration, for preventing and treating diseases in spinal cord, thyroid gland, ovary, prostate, renal gland, small intestine, heart, trachea, thymus, lymph node, muscular system and colon, for treating lipase deficiency in cystic fibrosis and pancreatitis, for treating undesirable clot formation such as myocardial infarction, during angioplasty and all surgical procedures that require decreased blood clot formation, for treating liver diseases, coagulation disorders and microbial diseases, for treating immune disorders, for treating inflammation and transplant rejection, for enhancing bone thickness and increasing bone density, for reducing the loss of essential ferroxidases, for suppressing apoptosis, and for regulating vascular smooth cell proliferation. They may also be used as vaccines.

Sequence 1887 BP; 259 A; 736 C; 577 G; 315 T; 0 other;
Query Match 20.2%; Score 634.6; DB 22; Length 1887;
Best Local Similarity 61.3%; Pred. No. 4.6e-124;
Matches 1079; Conservative 0; Mismatches 664; Indels 16; Gaps 3;

FT	XX	507	GTTTGCGGTGGTGGAGCCCTGCCCAAGTACTGTGTGTCGCCAGAAATCTGCTGAGTCACT	566
XX	DB	63	GTGAGCCACACCCAGCCCAATGTCCTCCCGCGCTGCCAGACACAGTCCGTCGCCCT	122
XX	QY	567	GGGACCTGTGCCCTTCCAAAGGGTGTCTTTGTACCCCTGATTTGACCGGGGAC	626
XX	DB	123	AAGCGTGTGTGCCAGGGGAGGCTCTGTGTGTCGCCACCTCGCTGACCGCGGGC	182
XX	QY	627	AGTGGAGCTGGCGCTGGCGGCAACTTTCATATCCACATCAGCCGCGAGACTTTCGCA	686
XX	DB	183	AGCGAGCTGGCGTGGCAGACAACTTTCATCGCTCCGTCGCGCGCGGAGCTGGCCAA	242
XX	QY	687	CATGACGGGCTGGTGGACTGACCTGTCAGGACACATCCAGCCACATCCAGCCCTT	746
XX	DB	243	CATGACAGGCTGCTGATCTGAGCTGTGCGGAAACATCCGCCACGTCGCTGCGG	302
XX	QY	747	TTCTTTTCGACCTGGAGCCCTCCGCTCCCTGCAATTTGACAGAAATCGTGCACG	806
XX	DB	303	CGCTTCGCGACTGGGGCCCTGGCTGCCCTGCACCTGGATGGAAACCGGTGACCTC	362
XX	QY	807	CGTGGGAGGACACCTCCGGGCGCTCACTCCAGCTGCAGACCTTATCGTGAACAA	866
XX	DB	363	ACTGGCGAGGGCAGCTGCGCGCTGGTCAACTTGGCCACCTCATCTCAGCAACA	422

Query Match 20.2%; Score 634.6; DB 22; Length 1887;
Best Local Similarity 61.3%; Pred. No. 4.6e-124;
Matches 1079; Conservative 0; Mismatches 664; Indels 16; Gaps 3;

AAH78215 standard; DNA; 1887 BP.
AAH78215;
26-NOV-2001 (first entry)
Nucleotide sequence of a human secreted polypeptide.
Human; secreted polypeptide; nervous disease; muscular disease; tumour; gastrointestinal ulceration; spinal cord disease; trachea disease; thyroid gland disease; ovary disease; prostate disease; heart disease; renal gland disease; small intestine disease; thymus disease; lymph node disease; muscular system disease; colon disease; lipase deficiency; cystic fibrosis; pancreatitis; clot formation; myocardial infarction; angioplasty; liver disease; coagulation disorder; microbial disease; immune disorder; inflammation; transplant rejection; bone thickness; bone density; ferroxidase loss; apoptosis; vascular smooth cell proliferation; vaccine; ss.
Homo sapiens.
Key Location/Qualifiers
1..1887 FT
CDS /*tag= a

1407 GGACCCAGCCCTTTATCCACTGGGTAGCCCCCGAATGACCCCTGGTAGGAATCCTC 1466
 963 GGACCCAGAGCCCGTGTGGTGTGACCCAGGGCCGGTGTAGGCAACTCAAG 1022
 1467 AAGGACCGTGTATGACAAAGGACCCCTGGACATTTTCATCACCACTCTCAGGACAG 1526
 1023 CCCTGCGCGCCCTTCCCAATGGGACCGTGGAGCTGTGGTACCAGCCGGGTATGG 1082
 1527 TGTGCTTCCACCTGCAATGCTCCCAATGTCGGAGAGGCAACCGCCATGTTGAGGT 1586
 1083 TGGCACTTCCACCTGCAATGCGGCCAATGACGCTGGAGGACACAGCTGCTGTGGAGCT 1142
 1587 CTCATGTCGACGTCACACCTCAGCAACAGCAGCAGCCGCA--CTGACCCCCCAAG 1644
 1143 GACTGTGGTCCCGCCACCTCTCAGTTCAGTTCAGCAACAGCAGCTGTGACCCCGCG 1202
 1645 TCCGCTCTCAGACATCACTGCTCCAGCAAGCAGCCGCGGAGGTGGAGGAGTGGG 1704
 1203 GGACGGGATCCTGATGCTCTCACCCACCTCCGCTCGCTTCTGCTTCCCAAGTGGC 1262
 1705 GGCGGAGAGCTCCCAAAAGCCCGGAAACCGGCTGCTGCTGCTGCTGCTGCTGCTG 1764
 1263 CGACACTGGGCCCCCTA-----CCGACCGTGGCTCCAGAGTACTGAGCAACGGG 1314
 1765 ACCTGGCCCTGTCAGTGTCTGTCAGCAAGTTCAGCAGCCCGGGGTGAAGATGACCA 1824
 1315 ACAGTGTCTTGTCCAGTGGCGGATCAGGGGCTATCCGGGCTATCCGCATGTACCA 1374
 1825 CTGAGTAACTGCTCTGACGATGAGTACTGATTTACAGGATGATCCAGCCTTCAAC 1884
 1375 ATCCAGTAAACAGCTCGGCTGATGACATCCTGCTTACAGGATGATCCCGGGGAG 1434
 1885 AAGCCCTTGTGGTCAACAACCTGGTTCAGGACTGGCTACGACTTGTGTGCTGGCC 1944
 1435 CGCTGCTTCTGTCAGCGGACTGGCGTTCAGGCGGACCTACGATCTGTGGTGTCTG 1494
 1945 ATGTGGATGACAGCCAGCAGCAGCTACCGCCCAACAACTCGTGGGCTGCCAGTTC 2004
 1495 GTGTATGAGGAGTGCACCGGGCTCACGGCCAGCGGCTGTGGGCTGGCCCGCTTC 1554
 2005 TTCCACCAAGGCTGACTACCCGAGTGCAGTCCATGCAAGCCAGATTTCTGGGGCACC 2064
 1555 TCCACCGAACCCTCGCTGGCCATGCGGGGCGCCACGCTCCCTTCTGGGGGCGCAG 1614
 2065 ATGATCTGCTATCGGGGATCATCTGCGCAGCTGCTGGCTTCTCATCGTCAATCCT 2124
 1615 ATGATCTGCTGCTGGGGGCTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1674
 2125 ATGCTGCGCTACAAGTCTGCAACCAACGAGGCCCCCAG-----CAAGATGCGAGG 2178
 1675 CTAAATGCGCTACAAGTGCACCGGCGCAGCCCCCGGCAAGGCAAGATTCGCCGCCC 1734
 2179 GTGAGCAATGTACTGACAGACCAACCGGCGCCAGCCCGCTCCCAAGCAGCCACCA 2238
 1735 GTTAGCAGGCTTGTCTCCAGACCAACCGGCGCCCTGGGCCCCCAGCCCGCCCGCC 1794
 2239 GCCGGGCCCCCGGCGAGG 2257
 1795 CCGCGCCGAGGCCCGCGG 1813

CC aberrant cellular adhesion, proliferation or differentiation. Specific.
 CC examples include melioidosis, juvenile AIDS, diabetes mellitus,
 CC rheumatoid arthritis, systemic lupus erythematosus, multiple sclerosis,
 CC Chron's disease, ulcerative colitis, asthma, anemia, and chronic active
 CC hepatitis. The present sequence represents the coding sequence of the
 CC human 33395 polypeptide.
 XX SQ Sequence 1887 BP; 257 A; 735 C; 579 G; 316 T; 0 other;
 Query Match 20.1%; Score 633; DB 22; Length 1887;
 Best Local Similarity 61.3%; Pred No 1e-123;
 Matches 1078; Conservative 0; Mismatches 665; Indels 16; Gaps 3;

507 GTTTCGCGTGTGACGCGCTGCCCAAGTACTGTGTGTCGCAAGATCTGTCTGACTACT 566
 63 GTACGACACACCCAGCCCATGTCCTCCCGCGCTGCCGCTCCAGACACAGTCTGCCCT 122
 567 GGGGACCTGTGCCCTTCAAGGGTGTCTTTGAPACCCCTGATATTGACCGGGGAC 626
 123 AAGCGTGTGTGCCAGGGGACAGCCCTCTGTTGTCGACCCCTGCTGGACCGCGGGC 182
 627 AGTGGAGCTGCGCTGGGGGCAACTTTCATTCATCCACATCAGCCGCGAGACTTTGCCAA 686
 183 AGCCGAGCTGCGCTGGCAGCAACTTTCATGCGCTCCGTCGCGCGCGGCACTGGCCAA 242
 687 CATGACGGGCTGTGACCTGACCCCTGTCAGCAACCAATCAGCCAGCATCCAGCCCTT 746
 243 CATGACAGGCTGTGATCTGACCTGTCGCGAACAACATCCGCGCAGTGGCTGCCGG 302
 747 TTCTTCTGAGCTCGAGGCTCCGCTCCCTGCAATTCAGCAAGTCCGCTGCCAAG 806
 303 CGCTTTCGCGACTTGGGGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 362
 807 CTTTGGGAGGACACCTCCCGGCTGTGTCAGCTTTCAGCTGACACTTATCGTGAACA 866
 363 ACTGGCGAGGCGCAGCTGGCGGCTGTGCAACTTTCGCGCCTCATCTCAGCAACA 422
 867 CCAGCTGGGCGCATCCAGATGAGGCTTTGAGGACTTCTGCTGACATTTGAGGAGTCT 926
 423 CCAGCTGGCAGCGCTGGCGGCGGCGCTGATGATGTTGCCGAGACACTGGAGACCT 482
 927 GGACTCTCTTACAAACCTCCATGCGCTGCGCTGGGATTCCTGCTGACATTTGAGGAGTCT 986
 483 CGACTCTCTTACAAACCTCCAGAGCTGCTGCGGAGGCTGCGGCGGCTGGGCGCA 542
 987 CCTCACCACTGAGCTGAGCAACACTGCTGATGATGATGATGATGATGATGATGATG 1046
 543 CGTCAACACTTGGGCTTCGACCAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 602
 1047 AGACTGCAAACTGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1106
 603 CGGCTGCAAACTGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 662
 1107 TGATCCCATTTTGGCGCTCCAGGCTTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1166
 663 CGACCCACTTCTTCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 722
 1167 CTTGCTCTTTAGTTTGGGGTAACTTCACTGCAATTTGAGCTTCTCTGCTGCTGCTG 1226
 723 CCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 782
 1227 GAGGCTGAGCGGAGGATGACTTGAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1286
 783 TCGCTGCGCGGAGGAGACTTCCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 842
 1287 CTACTTCTGGCATGCTGCTGAGGAGTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1346
 843 CTACTTTTGGCGGCTGGCGGAGGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 902
 1347 CACACAAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1406
 903 CTCACCACTTGGCTGTGCCCGAGTGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 962

RESULT 10
 AAI66985
 ID AAI66985 standard; cDNA; 2637 BP.
 XX
 AC AAI66985;
 XX
 DT 30-JAN-2002 (first entry)
 XX Human leucine-rich repeat (LRR) family member, 33395 cDNA sequence.
 DE LRR; leucine rich repeat; 33395; cytostatic; anti-HIV; antidiabetic;
 KW antiarthritic; neuroprotective; dermatological; immunosuppressive; ss;
 KW

antiinflammatory; antiasthmatic; antiulcer; antianaemic; hepatotropic.

KW Homo sapiens.
 OS Location/Qualifiers
 FH 1..434
 FT *tag= a
 FT 435..2321
 FT *tag= b
 FT /product= "33395"
 FT /note= "the coding sequence is specifically claimed"
 FT 2322..2637
 FT *tag= c
 XX WO200172827-A2.
 XX 04-OCT-2001.
 XX 23-MAR-2001; 2001WO-US09470.
 XX 24-MAR-2000; 2000US-191863P.
 XX (MILL-) MILLENNIUM PHARM INC.
 XX Glucksmann MA;
 XX WPI; 2001-626254/72.
 XX P-PSDB; AAG65805.
 XX New polypeptide for preventing or treating disorders associated with
 XX cellular adhesion, proliferation or differentiation, comprises
 XX polypeptide 33395, a member of the leucine rich repeat protein family
 XX
 XX Claim 1; Fig 1; 133pp; English.
 XX The invention provides an isolated nucleic acid encoding a polypeptide
 XX of the leucine rich repeat (LRR) family, designated 33395. The 33395
 XX polypeptide can be expressed by standard recombinant methodology. The
 XX 33395 polynucleotides and polypeptide can be used to prevent or treat
 XX disorders associated with 33395 expression, for example those involving
 XX aberrant cellular adhesion, proliferation or differentiation. Specific
 XX examples include melanoma, juvenile AIDS, diabetes mellitus,
 XX rheumatoid arthritis, systemic lupus erythematosus, multiple sclerosis,
 XX Chron's disease, ulcerative colitis, asthma, anemia, and chronic active
 XX hepatitis. The present sequence represents the cDNA encoding the human
 XX 33395 polypeptide.
 XX Sequence 2637 BP; 452 A; 976 C; 758 G; 451 T; 0 other;
 XX
 XX Query Match. 20.1%; Score 633; DB.22; Length 2637;
 XX Best Local Similarity 61.3%; Pred. No. 1.1e-123;
 XX Matches 1078; Conservative 0; Mismatches 665; Indels 16; Gaps 3;
 QY 507 GTTTCGGTGTGAGCGCTGCCAAGTACTGTGTCTCCAGAAFTCTCTGAGTCACT 566
 Db 497 GTCAGCCACACCCAGGCCATGTCCCGCGCTGCCAGACACAGTCTGCCCT 556
 QY 567 GGGACCTGTCCCTCAAGGGTGTCTTTGTACCCCTGATATTGACGGGGAC 626
 Db 557 AAGCGTGTGTGCCAGGGGAGGCTCTCTGTGTGCCACCCCTCGCTGGACCGCGGGC 616
 QY 627 AGTGAGCTGCGCTGGGGGAACTTTCATCATCCACATCAGCCGCGAGCTTTGCCAA 686
 Db 617 AGCCGAGCTGCGGTGGCAGACAACTTCATCGCCTCCGTGCGCGCGGACTGGCCAA 676
 QY 687 CATGACGGGGTGTGGACCTGACCTGTCCAGGAAACACATCAGCCCAATCCAGCCCTT 746
 Db 677 CATGACGGCCCTGCTGCTGATCTGAGCTGTGGGAAACACATCCGCGACAGTGGCTCCGG 736
 QY 747 TTCCCTTTCGACCTCGAGGCTCCGCTCCCTGTGATTTGACGGAATCGGCTGCCAAG 806
 Db 737 CGCCTTCGCGGACCTGCGGGCCCTCGCTCCCTGACCTGGATGGCAACCCGCTGACCTC 796

QY 807 CCTTGGGAGACACCTCCGGGGCTCGTCAACCTGCGAGCACCTTATCGTGAAACAACA 866
 Db 797 ACTGGCGAGGGCCAGCTGCGCGGCTGTTCAACTTGGCCACCTCATCTCAGCAACA 856
 QY 867 CCAGCTGGGGGCGATCGCAGATGAGGCTTTTGAGGACTTCCCTGCTGACATGGAGACT 926
 Db 857 CCAGCTGGCAGCGCTGGCGGCGGCGCTCGATGATGTTGCGGAGACACTGGAGACT 916
 QY 927 GGACCTCTCCATAACAACCTCCATGGCTGCGGTGGGACTCCGTGCCAGCATGGTCAA 986
 Db 917 CGACCTCTCTACAACAACCTCGAGAGTGCCTTGGGAGGCGCTGGCGCGCTGGCAA 976
 QY 987 CCTCCACAGCTGAGCCTGGACCAACACTGCTGTGATCACCTCCAACTGGCTGCGAAGCTGCC 1046
 Db 977 CGTCAACAGTGGGCTCGACCAACACTGCTGTGCTTGTGTGCCGCGCGCTTTTC 1036
 QY 1047 AGACCTCAGAAACTGGCCCGCTGGATCTCACCTCCAACTGGCTGCGAAGCTGCC 1106
 Db 1037 CGGCTGCACAAGCTGGCCCGCTGGACATGACCTCCAAACCGCTGACCAACTCCACC 1096
 QY 1107 TGATCCCATCTTTGGCCCGCTCCAGGCTTTGACAGCCACACACTTTGGCCCCACC 1166
 Db 1097 CGACCCACTTCTCCCGCTGCCCTGCTGCCCAGGCCCGGGGCTGCCCGCTCTG 1156
 QY 1167 TTGTCTTTAGTTTTGGGGTAAACCCACTTCACTGCAATTTGTGAGCTTCTCTGGCTGG 1226
 Db 1157 COTGGTGTGGCTTTGGCGGAAACCCCTGCACTGCAACTGCGAGCTGGTGTGGCTGG 1216
 QY 1227 GAGGCTCAGCGGAGCATGACCTGGAAACCTTGGCTCCCGAGGGGCGCTCAAGGGTGG 1286
 Db 1217 TGGCTGGCGGGAGGAGCAGCTCGAGGCTTGGCGGCTCCACCTGTCTGGGGGGCG 1276
 QY 1287 CTACTTTGGCATGTGCGTGGAGGAGTGTGTGCGAGCGGCTCTCATCCACGCA 1346
 Db 1277 CTACTTTGGCGGTGGGAGGAGGAGTGTGTGCGAGCCCGCTGGTGTACTCACCG 1336
 QY 1347 CACACAAGTTGTGTTTGGAGGGCAGCGCCACACTCAAGTGCAAAGCCATTGG 1406
 Db 1337 CTCACCACTCTGGTGTGCCCGAGTGGCGGCTGCTGCTGCTGCGGGCAGTGG 1396
 QY 1407 GGACCCAGCCCTTATCCACTGGTAGCCCCGATACCCCTGTGTAGGAACTCCTC 1466
 Db 1397 GGACCCAGAGCCCGCTGTGGTGTGCTACCCAGGGCGGGCTGTAGGAACTCAAG 1456
 QY 1467 AAGGACCGCTGTATGACAAATGGACCTGGACATTTTCATCACACATCTCAGSACAG 1526
 Db 1457 CCGTCCCGGCGCTTCCCAATGGAGCGCTGGAGCTGTGTACCGAGCCCGGTGATGG 1516
 QY 1527 TGGTCCCTTCACTGCTTGGCAATGTGCGGAGAGGCGCCAGCGGCGCATGGTGGAGT 1586
 Db 1517 TGGCATTTCACTGCTTGGGGCAATGCAGCTGGCGAGGCCACAGCTGTGTGGAGCT 1576
 QY 1587 CTCATGTCAGCTGCCACCTCAGCAACAGCAGCAGCCGCA--CTGCACCCCAAG 1644
 Db 1577 GACTGTGGTCCCCCACCCTCTCAGTAGCCAAACAGCAGCAGCTGTGACCCCCCGG 1636
 QY 1645 TCCCGCTCTCAGACATCACTGGCTCCAGAAAGACAGCCGGGAGGTGGAGGAGTGG 1704
 Db 1637 GGACGGGATCTGATGTCTTCCACCCACCTCCCTGCTGCTGCTGCTGCTGCTGCTG 1696
 QY 1705 GCGGAGAGCTCCCAAAAGCCCCCGAAAGGGGTGTGCTGTGTGTAAGTGCACACC 1764
 Db 1697 CGACACTGGGCCCTA-----CCGACCGTGGCGTCCAGGTGACTGAGCAGCGGGCC 1748
 QY 1765 ACCTCGGCCCTGTCAGTGTGTCAGCAAGTGCAGACCCCGGGTGAAGATGTACCAAG 1824
 Db 1749 ACAGTGTCTTGTCCAGTGGCGGATCAGCGGCTATCCCGGGATCCGCATGTACCA 1808
 QY 1825 CTGCACTACACTGCTGCGATGAGTACTGATTTACAGGATGATCCACGCTCCAC 1884
 Db 1809 ATCCAGTACAACAGCTCGGCTGTGATGACATCTCTGCTACAGGATGATCCCGGGAGAGC 1868

Novel proteins and polynucleotides of secreted proteins useful for treating various diseases e.g. rheumatoid arthritis, cancer, psoriasis, diabetic retinopathy, arteriosclerosis, ischemia or reperfusion injury

Claim 1; Page 124-127; 145pp; English.

The invention relates to human secreted polypeptides designated LP095, LP191, LP217, LP220, LP221, LP222, LP229, LP237 or LP238 and nucleic acid molecules encoding such polypeptides. Novel secreted proteins of the invention are used for treating diseases such as arteriosclerosis, Alzheimer's disease, diabetic retinopathy, psoriasis, pancreatitis, arteriosclerosis, rheumatoid arthritis, colorectal adenomas, severe combined immunodeficiency, ischaemia, carcinomas, haemolytic anaemia, reperfusion injury, neoplasms and cancer especially liver cancer. They are also used for wound healing. Polynucleotides of the invention can be used to generate transgenic animals or knock out animals. They turn, are useful in the development and screening of therapeutically useful reagents for use in the treatment of diseases associated with LP polypeptide associated activity. They are also used in gene therapy.

The present sequence is human LP220 secreted protein encoding cDNA. LP220 gene is located on chromosome 11q13.

PT	Novel proteins and polynucleotides of secreted proteins useful for	50	Query Match	19.8%	Score 622.6;	DB 24;	Length 2369;
PT	treating various diseases e.g. rheumatoid arthritis, cancer, psoriasis,	50	Best Local Similarity	62.3%	Pred. No. 1.6e-121;		
PT	diabetic retinopathy, arteriosclerosis, ischemia or reperfusion injury	50	Matches 1089;	Conservative	0;	Mismatches 609;	Indels 51;
PT		50					Gaps 5;
XX		50					
PS	Claim 1; Page 124-127; 145pp; English.	50					
XX		50					
CC	The invention relates to human secreted polypeptides designated LP095,	50					
CC	LP191, LP217, LP220, LP221, LP222, LP229, LP237 or LP238 and nucleic	50					
CC	acid molecules encoding such polypeptides. Novel secreted proteins of	50					
CC	the invention are used for treating diseases such as arteriosclerosis,	50					
CC	Alzheimer's disease, diabetic retinopathy, psoriasis, pancreatitis,	50					
CC	arteriosclerosis, rheumatoid arthritis, colorectal adenomas, severe	50					
CC	combined immunodeficiency, ischaemia, carcinomas, haemolytic anaemia,	50					
CC	reperfusion injury, neoplasms and cancer especially liver cancer. They	50					
CC	are also used for wound healing. Polynucleotides of the invention can	50					
CC	be used to generate transgenic animals or knock out animals. They	50					
CC	turn, are useful in the development and screening of therapeutically	50					
CC	useful reagents for use in the treatment of diseases associated with	50					
CC	LP polypeptide associated activity. They are also used in gene therapy.	50					
CC	The present sequence is human LP220 secreted protein encoding cDNA.	50					
CC	LP220 gene is located on chromosome 11q13.	50					
XX		50					
XX		50					
QQ	Sequence 2369 BP; 342 A; 855 C; 755 G; 417 T; 0 other;	50					

QY	520	GACGCTCCACCAAGTACTGTCGCCAGAACTGTCGTCGAGTCACTGGGGACCCTGTGC	579
DB	209	GCCGCTCCGCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	268
QY	580	CCCTCCAAAGGGGCTGCTTTGTATCCCTGATATTGACCGGGGACAGTGGAGCTGGC	639
DB	269	GCCACCCGAGGCTGCTTGTGTCGCGCCAAACGTGGACCGCGCCACAGTGGAGCTGGC	328
QY	640	CTGGGGCAACTTCATCCATCCAGCCCGCCAGGACTTTCACATGACGCGGGCTG	699
DB	329	CTGGTGTACAATTCATCCAGGCTTGGGGCCCTGACTTCCGACATGACGGACTG	388
QY	700	GTTGACCTGACCTTCCTCCAGGAACACCATGACGACATCCAGCCCTTTTCCTTCTGGAC	759
DB	389	GTTGACCTGACACTGCTCCGAAATGCCATCACCAGGACTTGGGGCCGCTTTGGGGAC	448
QY	760	CTCAGAGGCTCCGCTCCCTGCACTTTCAGAGCATCGGCTGCCAAGCTTTGGGGAGAC	819
DB	449	CTCAGAGGCTCCGCTCCCTGCACTTTCAGAGCATCGGCTGCCAAGCTTTGGGGAGAC	508
QY	820	ACCTCCGGGGCTGTTCAACCTGAGCAGCACCCTTTTCGTGAAACAACAGCTGGGGCGC	879
DB	509	AGCCTCCGGGGCCCGTCAATCTGAGCAGCACCCTTCCTCAGGGCAACAGCTGGGGCGC	568
QY	880	ATCGCAGATGAGCTTTTGAGGACTTCCTGCTGACATTTGGAGGATCTGGACCTCCCTAC	939
DB	569	ATCGCCCGGGAGCTTTCAGCAGCTTCTTAGAGAGCCTGGAGGACTGGACCTGTCCCTAC	628
QY	940	AACAACCTCATGGCTGCCGTGGGACTCCGCTGGCAGCAGCATGTTCAACCTCCACAGCTG	999
DB	629	AACAACCTCCGGCAGGTGCCCTGGCGGCACTGGGGCCATGCCTGCCCTGCACACCTTC	688
QY	1000	AGCCTGGACCAACAACCTGCTGGATCATATCGCCGAGGGACCTTTTCAGACCTGCAGAAA	1059
DB	689	AACCTGGACCAATAACCTTATTGACGCACTGCCCGAGGCGCTTCCGCGGCTGGCTAG	748
QY	1060	CTGGCCCGCTGATCTCACTCCAACTCGGCTGCAGAACTGCAGAGCTGCCCTCATCCCTTT	1119
DB	749	CTCTCCCGCTGGACCTCACTCCAAACCGCTGGCCAGCTGGCTCCGACCGCTTTTC	808
QY	1120	GCCGCTCCAGGCTTCGGCTTTGACAGCCACACCCTTTGGCCCCACCTTGTCTTCTTAGT	1179
DB	809	TCCTGGGGTGTAT-----GACAGGCTTCTCCGCGCCCTTCGGGCTGGCTGAGC	856

QY	1885	AAGGCTTCGTTCAACAACCTGTTGTCAGGACTGGCTACGACTTGTGTGCTGGCC	1944
DB	1869	CGCTCGTTCTCTGACGGACCTGGCGTCAAGCCGCGGACTACGATCTGTGGTCTCGCC	1928
QY	1945	ATGTGGGATGACACAGCCAGCACACTCAAGCCCAACAATCGTGGGCTGGCCAGTTC	2004
DB	1929	GTGTATAGGGCAGTGCACAGGGGCTCACGGCCACGGCGCTGTGGGCTGGCCCGCTTC	1988
QY	2005	TTACCAAGCTACTACCCGAGTCCAGTCCATGTCACAGCCAGATTTCTGGGGCACC	2064
DB	1989	TCCACCGAACCTGCGCTGGCGCATCGGGGGCCGACAGCTCCCTTCTGGGGGACG	2048
QY	2065	ATGATCCTGTCATCGGGGGCATCATCTGCGCCAGCTGTGGTCTTTCATCGTATCCTC	2124
DB	2049	ATGATCCTGTCATCGGGGGCATCATCTGCGCCAGCTGTGGTCTTTCATCGTATCCTC	2108
QY	2125	ATGGTCCGTACAAAGTCTGCAACCCAGAGGCCCCAG-----CAAGATGGCAGCGGCC	2178
DB	2109	CTAATCGCTACAAGGTGCACGGGGCCAGCCCGGCAAGGCAAGATTTCCCGGCC	2168
QY	2179	GTCAGCAATGCTACTCGCAGACCAACGGGCGCCAGCCACCGCTCCCAAGCAGCCACA	2238
DB	2169	GTTAGCAGGCTTGTCTCCAGACCAAGGGGCGCTGGGCGCCACAGCCCGCCCG	2228
QY	2239	GCCGGGCCCCCGGACG 2257	
DB	2229	CCCGCCCGGACCGCGG 2247	

RESULT 11			
ID	AAD38695	standard; cDNA; 2369 BP.	
XX	AAD38695;		
AC			
XX			
DI	23-SEP-2002	(first entry)	
DE	Human LP220 secreted protein encoding cDNA.		
XX	Human; secreted protein; arteriosclerosis; Alzheimer's disease; LP220;		
KW	diabetic retinopathy; severe combined immunodeficiency; pancreatitis;		
KW	rheumatoid arthritis; colorectal adenoma; haemolytic anaemia; cancer;		
KW	reperfusion injury; arteriosclerosis; wound healing; transgenic animal;		
KW	gene therapy; neoplasm; transgenic; psoriasis; ischaemia; carcinoma;		
XX	Chromosome 11q13; gene; ss.		
OS	Homo sapiens.		
XX			
PH	Key	Location/Qualifiers	
FT	167..2074		
CDS	/*tag= a		
FT	/product= "Human LP220 secreted protein"		
FT	167..214		
FT	/*tag= b		
FT	215..2071		
FT	/*tag= c		
FT	/product= "Mature human LP220 secreted protein"		
XX			
PN	WO200226801-A2.		
XX			
PD	04-APR-2002.		
XX			
PF	14-SEP-2001; 2001WO-US26026.		
XX			
PR	28-SEP-2000; 2000US-236088P.		
XX	(ELIL) LILLY & CO ELI.		
XX	Su EW, Wang H;		
XX	WPI; 2002-471259/50.		
DR	P-PSDB; .AAE23980.		
XX			

QY 1180 TTTGGGTTAACCCACTTCACTGCATTAATTTGTAGCTTCTGTGCTGGAGGCTCAGCGG 1239
 Db 857 TTTAGCGGAACCCCTGCCTGCACTGCATTAATTTGTAGCTTCTGTGCTGGAGGCTCAGCGG 916
 QY 1240 GACGATGACCTGAAACCTTGGCTCCCAGGCGCCTCAAGGTCGCTACTTTCTGGCAT 1299
 Db 917 CCGGACGACTTGAACCTGTCCTCCCTCCCGCGCCTGGCGCCTACTTTCTGGCA 976
 QY 1300 TTGGTGGGAGGAGTGGTGTGCGAGCCCTCTCATCCAGCCACACACACAAGTTG 1359
 Db 977 GTGCCGAGGGGAGTTCTCTGTGAGCGCCTCAATGCGCCACACAGCGAGCCTC 1036
 QY 1360 TTGGTTCGGAGGCCAGCCGACACCTAAGTCAAGAAAGCATTGGGACCCAGCC 1419
 Db 1037 TGGGTGTGAAAGGCGCGCCCTGCTCGGTCGCGGCGCCTGGGTGACCCCGCCT 1096
 QY 1420 TTTATCCACTGGGTAGCCGACTGACCGCTGAGGAACTCTCAAGGACCGCTGTC 1479
 Db 1097 ACCATGCACTGGTCTGTGAGCCGCTGTTGTTGGCAACTCTCCCGAGCCGGCT 1156
 QY 1480 TATGAAATGGCACCTTGACATCTTTCACACACATCTCAGGACAGTGGTCCCTAC 1539
 Db 1157 TTTCCCAAGGGACTTAGAGTGGGGTGACCGCGCTGGGACGCTGGGGCTACACC 1216
 QY 1540 TGCAATGTCGAATGTCCGGAGAGGCCACGGCCATGGTGGAGGTCTTCAATGTCAG 1599
 Db 1217 TGCAATGTCGAATGTCCGGAGAGGCCACGGCCATGGTGGAGGTCTTCAATGTCAG 1276
 QY 1600 CTGCCACACTCAGCAACAGCACAGCCGACTGCACCCCCCAAGTCCCGCCTTCAGAC 1659
 Db 1277 TTTGCCCCATGTGTTGGAACAGCATGCCGAGGGGCGCC-----CGGGCCCTCGGAC 1330
 QY 1660 ATCACTGGTCCAGCAACAGCACAGCCGAGGAGTGGAGGAGTGGGGCGGAGGCTCC 1719
 Db 1331 ATCGCGCCTCCGCTCGCACTGCTCCCGAGGTTGAGGGACCTGGAGTCTGAGCC- 1386
 QY 1720 AAAGCCCCCGGAAACGGGCTGTGCTGTGTGTGTAAGTACCAACAGCCCTTCGTC 1779
 Db 1387 -----AGCCCTGACGAGTACCGGAGTGGAGGAGTGGAGTCTGAGCC- 1429
 QY 1780 AAGTGGTGTGCAGCAAGTACCAACAGCCCGGAGTGTACAGCTGCAGTGCAGTACTGC 1839
 Db 1430 AGCTGGGTCCCAGGCGGCAAGCCAGTGTGGATTTCCAAATCCAGTACACAGC 1489
 QY 1840 TCTGAGATGAGTACTGATTTACAGATGATCCCAGCCTCCAAAGCCCTTCGTC 1899
 Db 1490 AGCGAAGATGAGACCTCATCTACCGGATTTCCAGCTCCAGCCACCATTTCTGCTG 1549
 QY 1900 AACAACTGGTGTGAGGGACTGGCTACGACTGTGTGCTGGCCATGTTGGGATGACACA 1959
 Db 1550 AAGCACTGTCCTCCCGGCGCTGACTATGACTGTGCTGTGCTGGCCTTGTCCACGGCGGT 1609
 QY 1960 G---CCACGACACTACGCGCCACCAACATCGTGGCTGCGCCCGTTCACCAAGGCT 2016
 Db 1610 GGGCCCTTGACTCATCGCCCAAGAGGCTGCTGGCTGTCCTTCTCCACAGTGGCG 1669
 QY 2017 GACTACCCGAGTGCAGTCCATGACAGCCAGATTTCTGGGCGGACCATGATCTGTC 2076
 Db 1670 GCTGCCCCGTGCCCCGCCCTGACAGCCAGTGTGGGGGACCCCTGACCTGGGCC 1729
 QY 2077 ATCGGGGATCATCGTGCCACCGTGTGCTTTCATCGTCTCCATGTTGGCTACTC 2136
 Db 1730 GTGGGGGGTGCTGTTGCTGCTTACTGTGCTTCAATTTGGCTTGTGCTTGGGGC 1789
 QY 2137 AAGTGTGACACAGAGCCCGCACAGAAATGGCAGGCGGCTGAGCAATGTGACTCG 2196
 Db 1790 CGGGGGCGGAAATGGCGCCCTCCCTCCAAAG-----CTCAGCCACGCTCCAGTCC 1840
 QY 2197 CAGACCAAGCGGCCAGCCAGCCGCTCCAAAGAGGGCACCAAGCGGCGCCCGGCGAG 2256
 Db 1841 CAGACCATTGAGGCCCCAGCCCCACCAAGGCCCCACCCCGGGGAGAGCCCCGGCCC 1900
 QY 2257 GGGCCCGCG 2265

Db 1901 CCGCGCAG 1909

RESULT 12
 ABK70010
 ID ABK70010 standard; DNA; 2855 BP.
 XX
 AC ABK70010;
 XX
 DT 15-JUL-2002 (first entry)
 XX
 DE cDNA encoding human Pro peptide #50.
 KW Human; ss; gene; PRO; secreted protein; transmembrane protein; genetic disorder; tumour; cancer.
 XX Homo sapiens.
 OS
 PN WO200224888-A2.
 XX
 PD 28-MAR-2002.
 XX
 PF 29-AUG-2001; 2001WO-US27099.
 XX
 PR 01-SEP-2000; 2000US-229896P.
 PR 05-SEP-2000; 2000US-230621P.
 PR 22-SEP-2000; 2000US-235147P.
 PR 10-NOV-2000; 2000WO-US30873.
 PR 12-JAN-2001; 2001US-261878P.
 PR 16-JAN-2001; 2001US-261910P.
 PR 16-JAN-2001; 2001US-261939P.
 PR 16-JAN-2001; 2001US-262150P.
 PR 25-JAN-2001; 2001US-264395P.
 PR 02-FEB-2001; 2001US-266421P.
 PR 09-FEB-2001; 2001US-267623P.
 PR 28-FEB-2001; 2001WO-US06520.
 PR 03-MAR-2001; 2001US-274399P.
 PR 03-APR-2001; 2001US-280982P.
 PR 04-APR-2001; 2001US-282129P.
 PR 04-APR-2001; 2001US-282199P.
 PR 09-MAY-2001; 2001US-290589P.
 PR 25-MAY-2001; 2001WO-US17092.
 PR 01-JUN-2001; 2001WO-US17800.
 PR 20-JUN-2001; 2001WO-US19692.
 PR 29-JUN-2001; 2001WO-US21066.
 PR 09-JUL-2001; 2001WO-US21735.
 XX
 PA (GETH) GENENTECH INC.
 XX
 PI Baker KP, Eaton DL, Filvaroff E, Goddard A, Grimaldi JC,
 PI Gurney AL, Smith V, Stephan J, Watanabe CK, Wood WI, Zhang Z;
 PI Fong S;
 XX
 DR WPI; 2002-362426/39.
 DR P-PSDB; ABG34079.
 XX
 PT New PRO polypeptides and polynucleotides encoding the polypeptides, useful in gene therapy, chromosome identification, tissue typing, or for genetic analysis of individuals with genetic disorders
 XX
 PS Claim 2; Figure 99; 218pp; English.
 CC This invention relates to the cDNA and protein sequences of novel secreted and transmembrane polypeptides PRO polypeptides. The invention also comprises a method for producing the proteins of the invention by recombinant means and antibodies specific for the protein of the invention. The antibody may be used for detecting the PRO proteins of the invention and may be used to modify their activity. CC polynucleotides may be used as hybridisation probes for a cDNA library to isolate the full-length PRO cDNA or to isolate other cDNAs, to construct hybridisation probes for mapping the gene which encodes that PRO and for genetic analysis of individuals with genetic disorders, in

assays to identify other proteins or molecules involved in binding reaction, to generate transgenic animals or knock-out animals which turn are useful in the development and screening of therapeutically useful reagents, for chromosome identification, and tissue typing. The PRO polypeptides are useful in gene therapy, and as molecular weight markers for protein electrophoresis purposes. The sequences may also be used to detect overexpression on PRO polypeptides in cancerous tumours and for screening for differentially expressed genes using microarray technology. The present sequence represents a cDNA encoding a human PRO protein of the invention.

Sequence 2855 BP; 552 A; 1007 C; 799 G; 497 T; 0 other;
 Query Match 19.8%; Score 621.8; DB 24; Length 2855;
 Best Local Similarity 61.3%; Pred. No. 2.5e-121;
 Matches 1078; Conservative 0; Mismatches 662; Indels 19; Gaps 4;

CC	507	GTTTCCGTTGGTCAGCCCTGCCCAAGTACTGTCTCCAGAACTCTCTGAGTCACT	566
CC	798	GTACGCCACACCAGCCCAATGTCCTCCCGCTGCCCTGCCAGACACAGTCCCTGCCCT	857
CC	567	GGGACCTGTGCCCTCCAAAGGGGTGCTTTTFTTACCCCTGATATTGACCCGGGAC	626
CC	858	AAGCTGTCTGCCAGGGGAGCCCTCCTGTTGTCGACCACTCCCTGGACCGCGGGC	917
CC	627	AGTGGAGCTGCGCTGGGGGCAACTTATATFCAATCATGACCCAGGACTTTGCCAA	686
CC	918	AGCCGAGCTGCGCTGGCAGACAACTTATCGCTCCGTGCGCCGCGGACCTGGCCAA	977
CC	687	CATGACGGGCTGGTGGACTGACCTGTCACAGAAACACCATCAGCCACATCCAGCCCTT	746
CC	978	CATGACAGGGCTGTGTGATCTGAGCCCTGTCGGGAAACCATCCGGCAGCTGGCTCGG	1037
CC	747	TTCCCTTCTGGACCTCGAGACCTCCCGCTCCCTGCAATTTGACGAAATCGGCTGCCAAG	806
CC	1038	CGCTTCGCGGACCTGGCGGCCCTGCTGTCCTGCACTGGATGGAAACCGGCTGACCTC	1097
CC	807	CTTTGGGAGACACCCTCGGGGCTGTTCACTCCAGCCTGACACTTATCGTGAACAACAA	866
CC	1098	ACTGGGCGAGGGCAGCTGGCGGCTGGTCAACTTGGCCCACTATCTCAGCAACAA	1157
CC	867	CCAGCTGGGCGGACCTCCAGATAGGCTTTTGGAGCTTCTGCTGACATTTGAGGATCT	926
CC	1158	CCAGCTGGCAGCGCTGGGGCGGCGCCCTGGATGATTTGCGCGAGACACTGGAGGACCT	1217
CC	927	GGACCTCTCTAACAACTTCAATGCGCTGCGGACTCCCTGGAGCCATGTTCAA	986
CC	1218	CGACCTCTCTAACAACTTCAATGCGCTGCGGACTCCCTGGAGCCCTGGCGCTGGGCAA	1277
CC	987	CTCCACAGCTGAGCCTGGACCAACACCTGCTGGATCACATCGCCGAGGGCACCTTTGC	1046
CC	1278	CGTCAACAGTTGGGCTTCGACCAACACTGCTGGCTTCTGTGCCC--GGCGTTTTTC	1334
CC	1047	AGACTGAGAAATGGCCCGCTGGATCTCACCTCCAATCGGTGAGAACTGTCGCCCC	1106
CC	1335	CGGCTGCAAGCTGGCCCGGCTGGATGATGACCTCCAAACCGCTGACCAATCCCAAC	1394
CC	1107	TGATCCCACTTTTCCCGCTCCAGGCTTCGGCTTTGACAGCCACACCTTTGCCCAACC	1166
CC	1395	CGACCCACTTCTCCCGCTGCCCCCTGCTCGCCAGGCGCCCGGGGCTGCCCGCTTGC	1454
CC	1167	CTTGTCCCTTTAGTTTTGGGGTAAACCACTTCACTGCAATTTGAGCTTCTTCTGGCTGG	1226
CC	1455	CTTGTGCTGGCTTTGGGGAAACCCCTGCATGCAACTGCGAGCTGGTGTGGCTGG	1514
CC	1227	GAGGCTGAGGGGACGATGACCTGAAACCTGTGGCTCCCGAGGGGCTCAAGGTTCC	1286
CC	1515	TCGCTTGGCGGGAGGACCTCGAGGCTTGGCGTCCCGCTTCCCGCTGCTGGCGGGC	1574
CC	1287	CTACTTCTGGCATGTGCGTGGAGGAGTTGTTGTCGAGCGGCTCTCATCACCCAGCA	1346
CC	1575	CTACTTCTGGCGGGTGGGGGAGGATTTGCTGCGAGGCCCGCTGCTGACTCACC	1634

QY	1347	CACACAAAGTTGCTGTTCTGAGGGCCAGGGCCACACTCAAGTGCAAAGCCATTGG	1406
DB	1635	CTCACCACTTGGCTGTCCCGCAGGTCCGGCGGTGCTCCCTGGCTGCGGCGAGTGG	1694
QY	1407	GGACCCAGCCCTTATCCACTGGGTAGCCCCCGATGACCCGCTGGTAGGAACTCCTC	1466
DB	1695	GGACCCAGCCCGTGTGGTTGGTGTACCCAGGGCCGGCTGCTAGGCAACTCAAG	1754
QY	1467	AAGGACCGCTGTATGACAAATGGACCTTGGACATCTTCATCACCACATCTCAGGACAG	1526
DB	1755	CCGTGCCCGGECCTTCCCAATGGGACCTGGAGCTGGTTCACCGAGCCGGTGTATGG	1814
QY	1527	TGTTGCCCTTACCTGCAATGCTCCAAATGTCCGAGAGAGCCACGGCCATGTGGAGGT	1586
DB	1815	TGGCATCTTACCTGCAATTCGGCCAAATGCACTGGCGAGGCCACACAGCTCTGTGGAGCT	1874
QY	1587	CTCAATCGTCAGTGCACACCTCAGCAACAGCCAGCCAGCCGCA--CTGCACCCCAAG	1644
DB	1875	GACTGTGGGTCCCGCCACCTCTCAGTAGTACCAACAGCACAGCTGTGACCCCGCGC	1934
QY	1645	TCCGCTCTCAGACATCACTGGCTCCAGCAAGACCCAGCCGGGAGGTGAGGCAFTGGG	1704
DB	1935	GGACGGGATCTGATGCTCTCACCCACCTCCGCTGCCTCTGCTTCTCCAAAGTGGC	1994
QY	1705	GGCGGAGAGCTCCCAAAAGCCCCCGGAAACCGGCTGTGCTGTGTGATGAGTGACCACC	1764
DB	1995	CGACACTGGGCCCCCTA-----CCGACCGTGGCTCCAGGATGATCCCGCGGAGAGC	2046
QY	1765	ACCTGGCCCTGTTCAAGTGTGTGTCAGCAAGTCTGACACCCCGGGTGAAGATGACCAAG	1824
DB	2047	ACAGTGTCTTTGTCAGTGGCCGGATCAGCGGCTATCCGGGCACTCCGATGTACCCAG	2106
QY	1825	CTGACAGTAACTGCTCTGACGATGAGTACTGATTTACAGGATGATCCAGCTCCNAAC	1884
DB	2107	ATCCAGTAAACAGCTCGGCTGATGACATCCTCTCTACAGGATGATCCCGCGGAGAGC	2166
QY	1885	AAGCCITCTGTGTCAACAACCTGGTTCAGGACTGGCTACGACTTGTGTGTGCTGGCC	1944
DB	2167	CGCTGTTCTGCTGACGGACCTGGCGTCAAGCGGACCTTACGATCTGTGTGCTGCTGCC	2226
QY	1945	ATGTTGGATGACACAGCCACGACACTCACGGCCACCAACATCGTGGGTGGCCCGAGTTC	2004
DB	2227	GTGTATGAGACAGCCACCGGGCTCACGGCCACCGGCTGTGGGTGGCCCGCTTC	2286
QY	2005	TTCCAAAGGTTGATACCCGAGTGCAGTCCATGACAGCCAGATTTCTGGCGGCGCAC	2064
DB	2287	TCCACCGAACTGCTGCTGGGCAATGCGGGCGGCGGACCTTCTTCTGGGCGGCGAGC	2346
QY	2065	ATGATCTGTTTATCGGGGCACTCATCTGGCCAGCTGCTGGTCTTTCATCTCATCTTC	2124
DB	2347	ATGATCATGCTGGCGGCTCATCTGAGCTCGTACTGTTTCTCATCTTCGTTGCTG	2406
QY	2125	ATGTTGGCTTACAAGTCTGCAACCAAGCCAGGGCCCCAG-----CAAGATGCGAGGGCC	2178
DB	2407	CTAATGGGTACAAGTGCACGGCGGCGAGCCCGCGGCAAGGCGCAAGATTCGCCGCT	2466
QY	2179	GTGCAAAATGTGTACTGCGAGACCAACCGGCGCCAGCCACCGCTCCCAAGCAGCGCA	2238
DB	2467	GTTAGCAGGGTTTGTCTCCAGACCAACCGGCGGCTTGGGCCCCCAAGCCAGCCGCGG	2526
QY	2239	GGCGGGGCCCGCGGCGAGG	2257
DB	2527	CCGCGCCCGAGCCCGCGG	2545

RESULT 13
 AAH78204
 ID AAH78204 standard; DNA; 2316 BP.
 XX
 AC AAH78204;
 DT 26-NOV-2001 (first entry)
 XX

DE Nucleotide sequence of a human secreted polypeptide.
 XX Human; secreted polypeptide; nervous disease; muscular disease; tumour;
 KW gastrointestinal ulceration; spinal cord disease; trachea disease;
 KW thyroid gland disease; ovary disease; prostate disease; heart disease;
 KW renal gland disease; small intestine disease; thymus disease;
 KW lymph node disease; muscular system disease; colon disease;
 KW lipase deficiency; cystic fibrosis; pancreatitis; clot formation;
 KW myocardial infarction; angioplasty; liver disease; coagulation disorder;
 KW microbial disease; immune disorder; inflammation; transplant rejection;
 KW bone thickness; bone density; ferroxidase loss; apoptosis;
 KW vascular smooth cell proliferation; vaccine; ss.
 XX Homo sapiens.
 OS
 XX
 FH Location/Qualifiers
 FT 1..2316
 FT CDS /*tag= a
 FT /*product= "secreted polypeptide"
 XX
 PN WO200166690-A2.
 XX
 XX 13-SEP-2001; 2001WO-US071143.
 XX
 XX 05-MAR-2001; 2001WO-US071143.
 XX
 XX 06-MAR-2000; 2000US-0187107.
 XX 13-MAR-2000; 2000US-0188916.
 XX 03-OCT-2000; 2000US-0236874.
 XX 03-OCT-2000; 2000US-0237846.
 XX
 XX (SMIK) SMITHKLINE BEECHAM CORP.
 XX (SMIK) SMITHKLINE BEECHAM PLC.
 PA Agarwal P, Murdoch PR, Rizvi SK, Smith RF, Xiang Z;
 XX
 XX WPI; 2001-570768/64.
 DR P-PSDB; AAG67512.
 XX
 XX Novel isolated secreted polypeptide useful for treating nervous and
 PT muscular diseases, gastrointestinal ulceration, coagulation and immune
 PT disorders, microbial diseases, inflammation and transplant rejection -
 XX
 XX Claim 2; Page 43-44; 102pp; English.
 XX
 XX The present sequence encodes a human secreted polypeptide. The
 CC secreted polypeptides and polynucleotides are useful for treating
 CC nervous and muscular diseases, for inhibiting tumour formation and
 CC metastasis, for treating gastrointestinal ulceration, for preventing
 CC and treating diseases in spinal cord, thyroid gland, ovary, prostate,
 CC renal gland, small intestine, heart, trachea, thymus, lymph node,
 CC muscular system and colon, for treating lipase deficiency in cystic
 CC fibrosis and pancreatitis, for treating undesirable clot formation
 CC such as myocardial infarction, during angioplasty and all surgical
 CC procedures that require decreased blood clot formation, for treating
 CC liver diseases, coagulation disorders and microbial diseases, for
 CC treating immune disorders, for treating inflammation and transplant
 CC rejection, for enhancing bone thickness and increasing bone density,
 CC for reducing the loss of essential ferroxidases, for suppressing
 CC apoptosis, and for regulating vascular smooth cell proliferation. They
 CC may also be used as vaccines.
 XX
 XX Sequence 2316 BP; 347 A; 836 C; 758 G; 375 T; 0 other;
 XX
 XX Query Match 19.7%; Score 618.8; DB 22; Length 2316;
 XX Best Local Similarity 60.7%; Pred. No. 1e-120;
 XX Matches 1129; Conservative 0; Mismatches 677; Indels 54; Gaps 5;
 XX
 XX 487 CTGCTAGCGTTTGGCATGGCTTTGGCGGTGCGAGCGCTGCCCGCCCAAGTACTGTCTGCG 546
 DB 61 CTGCTGCTCTGGCGGGGGGATCTGTTGGCCAGCCCTCTGGCGGGCCGCTGCACTCTGC 120
 XX
 XX 547 CAGAATCTGTAGTCACTGGGACCCTGTGCCCTCCAAGGGGTGCTCTTTGTACCC 606

Db 121 CAGAACGTGGCCCCACACTGACAAATGCTGTGGCCAAAGACCGGTTGCTCTTTGTGCGG 180
 QY 607 CCGTATATTACCGCGGACAGTGAGCTGCGCTGGCGGCAACTTTCATATCCACATC 666
 Db 181 CCGCGCATGACCGCGGGTGTGGCTGCGGCTACCGCAACTTTCATCGCGCCGCTG 240
 QY 667 ACCGCCAGGACTTTGCCAATACGCGGGTGTGGACCTGACCTGTCCAGGAACACC 726
 Db 241 CGCCGCCGAGACTTCCGCAACATACGAGCTGTGGTGCACCTCTCTCCCGGAACCC 300
 QY 727 ATCAGCCATCCAGCCCTTTTCCCTTTCTGGACCTCGAGAGCTCGGCTCCCTGATCTT 786
 Db 301 ATCGCCCAAGTGGCAGCTGGCGCCCTTCGCCGACCTGCGTGCCTCCGGGCCCTGCACTG 360
 QY 787 GACAGCAATCGGCTCCAAAGCCTTGGGGAGGACACCTCCCGGGGCTGTGCAACCTGCAG 846
 Db 361 GACAGCAACCGCTGGCGGAGTGGCGGGGACACAGCTCCGCGGCTGGGCAACTCCCG 420
 QY 847 CACCTTATCGTGAACAACAACAGCTGGGGGATCGCAGATGAGGCTTTTGGAGACTTC 906
 Db 421 CACCTGATCTTGGAAACAACAACAGATCGCGGGTGGAGTCCGGCGCCCTTTGACGCTTC 480
 QY 907 CTGCTGACATGGAGATCTGGACCTCTCTACACAACCTCCATGCCCTGCCCTGGGAC 966
 Db 481 CTGTCCACCGTGGAGGACCTGGATCTGTCTACAAACCTGGAGGCGCTTCGCGTGGAG 540
 QY 967 TCCGTGCCAGGATGTTCAACCTCCACAGCTGAGCTGGACCAACAACCTGTGGATCAC 1026
 Db 541 GCGTGGCCAGATGGTGAACCTAAACCTCACGCTGGACCAACAACCTCATGACCAAC 600
 QY 1027 ATCGCGAGGGCACCTTTGCAGACCTGCAGAAACTGGCCCGCTGGATCTCACTCCAAT 1086
 Db 601 ATCGCGAGGGGACCTTCTGTGAGCTTCAACAAGTGTCCCTGTGGACATGACCTCCAAC 660
 QY 1087 CGGCTGAGAGAGCTGCCCTGTATCCCATCTTTGGCCGCTCCAGGCTTCGGGTTTGAACA 1146
 Db 661 CGCCTGCATAAACTCCCGCCGAGGGCTTCTTCTGAGGTGCGAGG-----GCACC 711
 QY 1147 GCCACACCTTTGCCCCACCTTTGCTTTAGTTTTGGGGTAAACCACCTTCACTCCAAT 1206
 Db 712 GGGCCAAAGCCCCACCCCGTGCAGCTGACCTTCCGGGGCAACCCCTGCACTGCAAC 771
 QY 1207 TGTAGCTTCTGCTGGGAGGCTCGAGCGGGACGATGACCTGGAAACTGTGGCTCC 1266
 Db 772 TCGAGTGTCTTGGCTGCGGGCTGACCCCGAGGAGGACTTAGAGACCTGCGCCACG 831
 QY 1267 CCAGGGGCTCAAGGCTGCTTCTTGGCAATGCTGTGAGGAGAGTGTGTGCGAG 1326
 Db 832 CCGCAACACTCACCGACCGCTTCTTGGTCCATCCCGAGGAGGAGTTCCTGTGTGAG 891
 QY 1327 CCGCTCTCATCACCACGACACAC---ACACAAGTTGCTGGTTCCTGAGGCGCGGGCC 1383
 Db 892 CCCCCTGTATCACACGGCAGCGGGGGCCCGGGCTTGTGTGAGAGCCAGGGCGGTG 951
 QY 1384 ACACCAAGTCAAAAGCCATTTGGGACCCACCCCTTATCCACTGGGTAGCCCCCGAT 1443
 Db 952 AGCCTGCGTGGCGAGCGGTGGTGCACCCGAGCGGTGGTGCACCTGGGTGGCACCTGAT 1011
 QY 1444 GACCGCTGTGAGGAACTCTCAAGGACCGCTGTCTATGACAATGGCAACCTTGACATC 1503
 Db 1012 GGGCGCTGTGGGAACTCCAGCGGACCCCGGGTCCGGGGGAGCGGACGCTGGATGTG 1071
 QY 1504 TTCATCACACATCTCAGGACAGTGGTGCCTTCACTGCATGCTGCCAATCTCCCGGA 1563
 Db 1072 ACCATCACACCTTGGGACAGTGGCACCTTCACTTGTATCGCTTCACTCAATGCTGGG 1131
 QY 1564 GAGGCCAGGGCATGGTGGAGTCTCCATCGTCCAGCTGCCACACCTCAGCAACAGCACCC 1623
 Db 1132 GAAGGACGGCCCGTGGAGGTGCGTGGTACCTCTGCTC-----TGATGGCACCC 1185
 QY 1624 AGCCGCACTGCACCCCAAGTCCCGCTCTCAGACATCACTGGGTCCAGCAAGACCAAGC 1683

mat_peptide 55:1653
 /*tag= C
 /product= "Mature ZLRR8 protein"

MO200202604-A2.
 10-JAN-2002.
 02-JUL-2001; 2001WO-US20999.
 30-JUN-2000; 2000US-215446P.
 (ZYMO) ZYMOGENETICS INC.
 Thayer EC, Sheppard PO, Presnell SR;
 WPI; 2002-154725/20.
 P-PSDB; AAE17484.

New leucine-rich repeat proteins and polynucleotides, useful for diagnosing and treating disorders related to abnormal cell growth e.g. retinoblastoma, renal cell, endometrial adenocarcinoma, leukemia, kidney and lung tumors

Claim 19; Page 64-67; 82pp; English.

The invention relates to an isolated polypeptide comprising leucine-rich repeat proteins Zlrr7, Zlrr8, Zlrr9 and spliced variants of Zlrr8, Zlrr9. Zlrr7, Zlrr8, and Zlrr9 proteins are useful in directing the secretion of proteins of interest from a host cell and to monitor the secretion of proteins in general from cells and tissues. The Zlrr DNA and proteins are useful in diagnosing and treating disorders related to abnormal cell growth, including retinoblastoma, renal cell adenocarcinoma, endometrial adenocarcinoma, glioblastoma, neuroblastoma, B-cell lymphotic leukemia, kidney tumours, germ cell tumours, lung large cell carcinoma, mammary, colon adenocarcinoma, genitourinary tract transitional cell tumours, rhabdomyosarcoma, lung tumour, bladder tumour, oesophagus, pancreas and prostate adenocarcinoma. Zlrr protein is useful for identifying agonists and antagonists of the polypeptide, for drug design, to screen for cell metabolism effecting receptors, for analysis of cell phenotype, and as animal feed supplement and cell culture components. Zlrr DNA is also useful in gene therapy. The present sequence is human Zlrr8 DNA.

Zlrr8 gene is located on chromosome 11q13.

Sequence 1653 BP; 223 A; 617 C; 526 G; 287 T; 0 other;

Query Match 19.3%; Score 608; DB 24; Length 1653;
 Best Local Similarity, 63.7%; Pred. No. 1.8e-118;
 Matches 1025; Conservative 0; Mismatches 540; Indels 43; Gaps 5;

520 GACGGCTGCCCAAGTACTGTCTGCGCAGAAATCTGTCTGAGTCACTGGGGACCCCTGTGC 579
 43 GCGGCTGCCGCTGCCCTGCGTCTGCCAGAAACCTGTCCGAGTCTCCAGACCCCTCTGT 102

580 CCTTCAAGGGTCTCTTTTACCCCTGTATTTGACCGGGGACACAGTGGAGTGGCG 639
 103 GCCACCGAGGGCTGTCTTTTGTCCGCCCAACCTGGAGCCGGCCACAGTGGAGTGGCG 162

640 CTGGGGGCAACTTCAATCATCCATCAGCCCGCCAGGACTTTGCCAAATGACGGGGCTG 699
 163 CTGGCTGAGAACTTCACTCCAGCCCTGGGGCCCTGACTCCGCAACATGACGGACTG 222

700 GTGGACCTGACCTGTCCAGGAAACACCATCAGCCATCAGCCCTTTTCCCTTTCTGGAC 759
 223 GTGGACCTGACCTGTCTCGCAATGCCATCACCCGCAATTTGGGGCCCGCCCTTTGGGAC 282

760 CTCGAGGCTCCGCTCCCTGCACTTTGACAGCAATCGGCTGCCAAGCCCTTGGGGAGGAC 819
 283 CTCGAGGCTCCGCTCCCTGCACTTTGACAGCAATCGGCTGCCAAGCCCTTGGGGAGGAC 342

820 ACCCTCCGGGGCTGGTCAACCTGACAGCACTTATCGTGAACAACAACACAGCTGGGGCC 879
 343 AGCCTCCGGGGCTGGTCAACCTGACAGCACTTATCGTGAACAACAACACAGCTGGGGCC 402

Db 1186 CGCCGGCTGCCCGCCCTCTACAGCCCGGCTCTCTGACATCGCCACCGCCGGC 1245
 Qy 1684 CGGGAGGTGAGGAGTGGGGGCGGAGCCCTCCAAAAGCCCGGAAAGGGGCTGTG 1743
 Db 1246 AGACAGGTGCCAACAGTTCGGGTGAGCGTC-----GGCTC 1284
 Qy 1744 CTTGTGCTGAAGTACACACACCTCGGCCCTGGTCAAGTGTCTGTCAGCAAGTCAGCA 1803
 Db 1285 GTGGAGCCGAGCTCACCTCGAATCCCGTCTATCCCGTGGCCAGCCAGAGGCTGTG 1344
 Qy 1804 CCCCGGTGAAGATGATACAGCTGACAGTCAACTGCTCTGACGATGAGGTAAGTATTTAC 1863
 Db 1345 CCGGMAATAGGATGATACAGGTTCACTACACAGTTCGGTTGATGACTCCCTGCTTAC 1404
 Qy 1864 AGGATGATCCAGCTCCAAAGGCTTCTGGTCAACAACCTGGTGTGACGGACTGGC 1923
 Db 1405 AGGATGATCCCGTCCACCACTTCCCTGGTGAATGACCTGGCGGGCCGCTGCC 1464
 Qy 1924 TAGCACTTGTGTGCTGGCCATGTTGGATGACACAGCCACGACTCACGGCCACCAAC 1983
 Db 1465 TAGCACTTGTGTGCTGGCCATGTTGGATGACACAGCCACGACTCACGGCCACCAAC 1524
 Qy 1984 ATCGTGGCTGCCCGCCAGTCTTCCAAAGGCTGACTACCCGCAAGTCCATGTCAC 2043
 Db 1525 GTGTGGCTGTGTACAGTTCACACACCGCTGGGATCCCGGCGCCCTGGCTGAGG 1584
 Qy 2044 AGCCAGATTCGGGGCCACCATGCTGGTCACTCGGGGCAATCATCGTGGCCAGCTG 2103
 Db 1585 GCCCATTTCTGGGGCCACCATGATCGCCATCGGGGGCTCATCTGCTGCGCTCGGT 1644
 Qy 2104 CTGGTTCATCTCATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2162
 Db 1645 CTGCTTTCATCTCTCATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1704
 Qy 2163 -----CRAGATGGCAGCGGCGCTGAGCAATGTACTGCGAGCAACAGCGC 2208
 Db 1705 CGGCTCAAGGGCTCCAGGTGCTCCCGCGGTGAGCCACGCTGTGCTCGCAGCAACGGC 1764
 Qy 2209 GCCAGCCCGCTCCAAAGGCGCCAGCCAGCCGGGGCCCGCCCGGGCCCGCGGAG 2268
 Db 1765 GCAGGCACAGCGCGGCGCAGCGCCCGCCCTGCGCGCCCGCCAGCACTACGAGGCGCTG 1824
 Qy 2269 GTGTGCTGCGCAACAGCTCTGACTTCCAGCCGCAAGCTGGCCCGCCCGCCAGTACTCC 2328
 Db 1825 CGGAGGTGAGTCCAGGCTGCCCCCGCGCTGCGCGTCCGAGCCCAAGGCTGAGGCGCC 1884

RESULT 14
 AAD28124
 ID AAD28124 standard; DNA; 1653 BP.
 XX AAD28124;
 AC AAD28124;
 XX
 DT 22-APR-2002 (first entry)
 XX
 DE Human leucine-rich repeat-8 (ZLRR8) DNA #2.
 XX
 KW Human; leucine-rich repeat-8; ZLRR8; cytostatic; gene therapy; leukaemia;
 KW endometrial adenocarcinoma; renal cell; colon; prostate; retinoblastoma;
 KW cell growth disorder; glioblastoma; neuroblastoma; kidney tumour; lung;
 KW mammary gland; germ cell; bladder; oesophagus; pancreas; animal feed;
 KW rhabdomyosarcoma; genitourinary tract; chromosome 11q13; ds.
 XX
 OS Homo sapiens.
 XX
 FH Location/Qualifiers
 FH 1..1653
 CDS /*tag= a
 FT /product= "ZLRR8 protein"
 FT /note= "CDS does not include stop codon"
 FT /partial
 FT 1..54
 FT /*tag= b
 FT sig_peptide

880 ATCCGAGATAGGCTTTGAGACTTCTGCTGACATTTGGAGGATTTGGACCTTCTCTAC 939
 403 ATCCGCGGGAGCCTTTCGACACTTCTAGAGAGCTTGGAGCCTTGGACCTTCTCTAC 462
 940 AACACCTCCATGGCTCCGCTGGGACTCCGCTGGACGATGGTCAACCTCCACAGCTG 999
 463 AACAACTCCGCGAGGCTTGGCCGCGCATCGGCGCCATGCCTGCCCTGCACACCTTC 522
 1000 AGCTGGACCAACCTGCTGGATCACATCGCGAGGGCACCTTTCAGACCTGCAGAAA 1059
 523 AACCTGGACCAATTAATTCAGCAGCTGCCCCAGGGCCCTTTCGCCAGCTCGGTGAG 582
 1060 TTGGCCGCTCGATTCACCTCCAAATCGGCTGCAGAAAGTCCGCCCTGATCCCACTTTT 1119
 583 CTCTCCGCTGGACTCACTCCAAACCGCTGGCCAGCCTGGCTCCGGACCCGCTTTTC 642
 1120 GCCCGCTCCAGGCTTCGGCTTTCAGAGCCACCTTTGCCCCACCTTTGCTTTAGT 1179
 643 TCTCGTGGCGGTAT-----GCAGAGGCTCTCCCGCCCTGCTGGTGGAGC 690
 1180 TTTGGGGTAAACCCACTTCACTGCAATTTGTGAGTCTCTGCTGGGAGGCTCGAGCGG 1239
 691 TTTAGGGGAACCCCTGACTGCAACTGTGAGTCTGTGCTGGGGGGCTGGCGGG 750
 1240 GACGATGACCTGGAAACCTTGGCTCCCGAGGGGCTCAAGGGTCTTCTTGGCAT 1299
 751 CCGGAGACTTGGAAACGCTGCCCTCCCGCCGCGCTGGCCGCGCTTCTTGGGCA 810
 1300 GTGCTGAGGAGGTTGTGTGGAGCCGCTCTCATCCAGCAGCACACACAGTTG 1359
 811 GTGCCGAGGGGAGTTCCTGTGTGAGCCGCTCAITGGCCGCGCACAGGCGCTC 870
 1360 CTGGTTCTGAGGGCCAGCGCCACACTCAAGTCAAGCAAGCATTGGGAGCCACCGCC 1419
 871 TGGTGTGGAAGCCAGCGGGCCAGCTTCCGTTCCGGTCCCGGGCCCTGGGTACCCCGCC 930
 1420 TTATTCACCTGGGTAGCCCGGATGACCGCTGGTAGGAACTCTCAAGGACCGGCTGTC 1479
 931 ACCATGCACTGGTGGTCTGACGACCGGTTGGTTGGCAACTCTCCCGAGCCCGGCT 990
 1480 TATGACATGGACCTTGACATCTTATCACCACATCTCAGGACAGTGGCTTCCACC 1539
 991 TTTCCCAACGGACCTTAGAGATTGGGGTGTACCGGCTGGGGAGCGCTGGGGGTACACC 1050
 1540 TGCATTGCTCAATGCTCCGAGAGGCGAGCCATGTTGGAGTCTCCATGCTCCAG 1599
 1051 TGCATGCCACCAACCTGCTGTGAGGCCACAGCCCGAGTAGAAGTGGGGTGTGGCC 1110
 1600 CTGCCACACTCAGCAACAGCAGCCGCACTGCAACCCCGGCTCCCGCTCTCAGAC 1659
 1111 TTGCCCATGGTGGAAACAGCAGTGCAGGGGGCGCC-----CGGGCCCTGGGAC 1164
 1660 ATCAGTGGCTCCAGCAAGACCGGGAGGTGGAGGAGTGGGGGGGAGAGCTCC 1719
 1165 ATCCGCGCTCCGCTCGCACTGCTGCGAGGCTGGGGAGCGTGGAGTCTGAGCC- 1220
 1720 AAAAGCCCCCGGAAAGGCTGTGCTGTTGCTGAAAGTGCACACCACTCCGCTCCGCTG 1779
 1221 -----AGCCGTGAGGTTGACGGAGGTGACCGCCACCTCAGGGCTGGG 1263
 1780 AAGTGGTCTGTGAGCAAGTCAAGCCCGGGTGAAGTGTACAGTGCAGTCAACTGC 1839
 1264 AGTGGGGTCCCGGGGGCCAGCCGACCCAGTGTGGATGTCCAAATCCAGTACACAC 1323
 1840 TCTGAGATGAGTGTATTTACAGATGATCCAGCCTCCAAAGGCGCTTGTGGTGC 1899
 1324 AGCGAAGTGAAGCCCTATCTACCGGATTTGCCAGCCTCCAGCCACCTTCTGCTG 1383
 1900 AACAACTGGTGTGAGGACTGGTACGACTTGTGTGCTGCGCATGGGATGACACA 1959
 1384 AAGCACTGTCCCGGGCGCTGACTATGACTTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 1443

1960 G---CCAGACACTCCGCGCACCAACATCTGTTGGCTGGCCCGCTTCTTACCAAGGCT 2016
 1444 GGGCCCTTGACCTCAGCGCCACCAAGGCTGTGGCTGTGGCCATTTCTCCACGCTGGC 1503
 2017 GACTACCCGCGAGTGCAGTCCATGCACAGCCAGATTTCTGGGGGACCAATGATCTGTG 2076
 1504 GCCTCGCCCTGTGCACGCTCGAGCCACAGTGTGGCGGAGCCCTGACCGTGGCC 1563
 2077 ATCGGGGGCATCATCTGTTGGCCACGC-TGCTGGTCTTTCATGTCATCT 2123
 1564 GTGGGGGGTGTGCTGTGGCTGCGCTTACTGGTCTTCACTGTGGCCCT 1611

RESULT 15
 AAD28125
 ID AAD28125 standard; DNA; 1653 BP.
 XX AAD28125;
 XX DT 22-APR-2002 (first entry)
 XX Human leucine-rich repeat-8 (ZLRR8) degenerate DNA #2.
 DE Human; leucine-rich repeat-8; ZLRR8; cytostatic; gene therapy; leukaemia;
 KW endometrial adenocarcinoma; renal cell; colon; prostate; retinoblastoma;
 KW cell growth disorder; glioblastoma; neuroblastoma; kidney tumour; lung;
 KW mammary gland; germ cell; bladder; oesophagus; pancreas; animal feed;
 KW rhabdomyosarcoma; genitourinary tract; chromosome 11q13; ds.
 OS Homo sapiens.
 XX WO200202604-A2.
 XX 10-JAN-2002.
 XX 02-JUL-2001; 2001WO-US20999.
 XX 30-JUN-2000; 2000US-215446P.
 XX (ZYMO) ZYMOGENETICS INC.
 XX Thayer EC, Sheppard PO, Presnell SR;
 XX WPI; 2002-154725/20.
 XX New leucine-rich repeat proteins and polynucleotides, useful for
 XX diagnosing and treating disorders related to abnormal cell growth e.g.
 XX retinoblastoma, renal cell, endometrial adenocarcinoma, leukemia,
 XX kidney and lung tumors
 XX Disclosure; Page 69-70; 82pp; English.
 XX The invention relates to an isolated polypeptide comprising leucine-rich
 XX repeat proteins Zlrr7, Zlrr8, Zlrr9 and spliced variants of Zlrr8, Zlrr9.
 XX Zlrr7, Zlrr8, and Zlrr9 proteins are useful in directing the secretion of
 XX proteins of interest from a host cell and to monitor the secretion of
 XX proteins in general from cells and tissues. The Zlrr DNA and proteins
 XX are useful in diagnosing and treating disorders related to abnormal cell
 XX growth, including retinoblastoma, renal cell adenocarcinoma, endometrial
 XX adenocarcinoma, glioblastoma, neuroblastoma, B-cell lymphocytic leukemia,
 XX kidney tumours, germ cell tumours, lung large cell carcinoma, mammary,
 XX colon adenocarcinoma, genitourinary tract transitional cell tumours,
 XX rhabdomyosarcoma, lung tumour, bladder tumour, oesophagus, pancreas
 XX and prostate adenocarcinoma. Zlrr protein is useful for identifying
 XX agonists and antagonists of the polypeptide, for drug design, to screen
 XX for cell metabolism effecting receptors, for analysis of cell phenotype,
 XX and as animal feed supplement and cell culture components. Zlrr DNA is
 XX also useful in gene therapy. The present sequence is human Zlrr8
 XX degenerate DNA. Zlrr8 gene is located on chromosome 11q13.
 XX Sequence 1653 BP; 179 A; 217 C; 329 G; 195 T; 733 other;
 SQ Query Match 17.6%; Score 554.6; DB 24; Length 1653;

Best Local Similarity 40.1%; Pred. No. 3.2e-107;
 Matches 643; Conservative 276; Mismatches 642; Indels 42; Gaps 4;

Db 1043 GNTAYACNTGYATHGNCVACNAAYCCNCGNGGARGCNCACNCGNMGNTNGARYTMMNG 1102
 Qy 1592 TCGTCCAGCTGCCACACCTCCAGCAACAGCACAGCCGCACTGCACCCCAAGTCGCCGCC 1651
 Db 1103 TNYTNGCNYTNCNCAYGGNGNAA{YWSNWS-----NGCNGARGGNGMGNCCNGENC 1156
 Qy 1652 TCTCAGACATCACTGGCTCCAGCAAGACCAGCCGGGAGGTGGAGGCACTGGGGGGGG 1711
 Db 1157 CNWSNGAYATHGNCNWSNWCNACNCGNCGNARGGNGARGVNACNYTNGARWSNG 1216
 Qy 1712 AGCCTCCAAAAGCCCGGAAACGGGCTGTCTGTTGTCTGAAGTGACCCACCCCTCG 1771
 Db 1217 ARCC-----NGCNGTNCARGTNCNCGNARGTNCNCGNACNWSNG 1255
 Qy 1772 CCCTGGTCAAGTGGTCTGTCCAGGAAGTCAGCACCCCGGGTGAAGATGTACCAGCTGCAGT 1831
 Db 1256 GNYTNGTWSNTGGGNCNCGNCGNARGTNCNCGNARGTNCNCGNACNWSNG 1315
 Qy 1832 ACAACTGCTCTGACGATGAGTACTGATTTACAGGATGATCCAGCCCTCCCAACAGGCTC 1891
 Db 1316 AYAA{YWSNWSNGARGAYGARACNYTNAHTAYMGNATHGTNCCNCGNWSNWSNCA{Y 1375
 Qy 1892 TCGTGGTCAACAACCTGGTGTCCAGGACTGGCTACGACTGTGTGTCTGGCCATGTGGG 1951
 Db 1376 TYYTNYTNAARCA{YTYTNGTNCNCGNCGNARGTAYGAYTNYTNYTNYTNGCNYTNSNC 1435
 Qy 1952 ATGACACAGCCAGCACACA---CTCACGGCCACCAACATCGTGGGCTGCAGTCTCTTCA 2008
 Db 1436 CNGCNGCNGNCGNWSNGAYYTNCNCGNACNWSNWTNYTNGGNTGYCNCAYTTYWSNA 1495
 Qy 2009 CCAAGGCTGACTACCCGAGTCCAGTCCATGCACAGCCAGATTCTGGCCGACCATGA 2068
 Db 1496 CNYTNCNCGNWSNCCNYTNYTGYCAYCNYTNCARGCNCAYTNYTNGGNGNACNYTNA 1555
 Qy 2069 TCCCTGGTCACTGGGGGCATCATCGGGCCAGCTGCCTGCTT 2111
 Db 1556 CNGTNGCNGTNGGNGGNTNYTNGTNGCNGNYTNCNCGNWT 1598

Search completed: June 8, 2003, 15:46:32
 Job time : 698 secs

Qy 512 CCGTGGTCGACCCCTGCCCAAGTACTGTGTGCTGCAGAACTCTGCTGAGTCACTGGGGA 571
 Db 35 SNGNGCNGCNGCNTGYCCNYTNCNTGYTGTNGYCARAA{Y{Y{WSNWSNGARWSN{Y{WSNA 94
 Qy 572 CCCTGTGCCCTCAAAGGGCTGCTTTGTACCCCTGATATGACCCGGGACAGTGG 631
 Db 95 CN{Y{NGY{CNCAY{MNG{NGY{NY{T{Y{GT{NCC{NAA{Y{GT{NGAY{MNG{NAC{NGT{NG 154
 Qy 632 AGCTGCCCTGGCGGCAACTTCATCATCCATCAGCCCGCCAGGACTTTCACACATGA 691
 Db 155 AR{Y{MNG{Y{NG{CNGAY{AA{Y{Y{TH{CARG{CNY{TNG{NCC{NCGN{CNGAY{T{Y{MNG{AA{Y{TA 214
 Qy 692 CGGGCTGTGGACCTGACCTGTCCAGGAACACCATCAGCCACATCCAGCCCTTTTCT 751
 Db 215 CNGGNYTNGTNGAYTNCN{Y{MNG{NAA{Y{C{NATH{C{MGNATH{GNG{CNGM{NGCNT 274
 Qy 752 TCTGTGACCTCGAGACCTCCGCTCCCTCATCTTGCAGACAAATGGTGGCCAAAGCCTTG 811
 Db 275 TYGNGAYTNGARWSNYTMMNGNCTNAA{Y{Y{TNGAY{GNAA{Y{MNG{Y{NGTNGAR{Y{TNG 334
 Qy 812 GGGAGACACCTCCGGGCTGGTCAA{CTGCAGCACCTTATCGTGAACAACAACACAGC 871
 Db 335 GNACNGNWSNYTMMNGNCGNTNAA{Y{Y{TNCARCA{Y{Y{TNA{H{Y{TWSN{GNGNAA{Y{CARY 394
 Qy 872 TGGGCGCTCGCAGATGAGCCTTTGAGGACTTCCTGCTGACATGGAGGATCTGGACC 931
 Db 395 TNGNMGNATHGNCNCGNCGNTY{Y{GAY{T{Y{Y{TNGARWSNYTNGAR{Y{Y{NGAY{Y 454
 Qy 932 TCTCTTAACAACCTCACTGGCTCGGCTGGGACTCCGTGCGACCATGGTCAA{CCTCC 991
 Db 454 TWSNTAYAA{Y{Y{MNG{CARGTNCNTGGGNGNATHGNGNCGNATGCCNCGNYTNC 514
 Qy 992 ACCAGCTGACCTGGACACAACCTGCTGATCACATCGCGGAGGCCACTTTGCAGACC 1051
 Db 514 AYACNYTNAAYTYNGAYCA{Y{AA{Y{Y{TNA{H{Y{GAY{CNY{TNCNCGNCGNCTY{GCNCARY 574
 Qy 1052 TGCAGAACTGGCCGCTGGATCTACCTCCAATGGTGCAGAAAGTCCGCCCTGTGATC 1111
 Db 574 TNGNCARYTNSNWSNYTNGAYYTNCNWSNAA{Y{MNG{Y{NGNACNYTNGCNCNGAYC 634
 Qy 1112 CCATCTTTGCCGCTCCAGGCTTCGGCTTGTGACAGCCACACCTTTGCCACCCTTGT 1171
 Db 635 CNYTNYTWSNMGNGNMGAY-----GCNGARGCNCNWSNCCNCGNCCNYTNG 682
 Qy 1172 CCTTAGTTTGGGGTAA{C{CACTTCACTGCAATTTGAGCTTCTGCTGGCTGGAGGC 1231
 Db 683 TNYTWSNTY{WSN{GNA{Y{C{NY{TNCAY{TGYAA{Y{GYARY{TNY{TGGY{TNNM{NGNY 742
 Qy 1232 TCGAGCGGACGATGACCTGAA{A{CCTGTGGCTCCCAGGGGGCCTCAAGGTCGCTACT 1291
 Db 743 TNGCNGCNGAYGAYTNGARACNTGYCNCNCCNCGN{Y{NGCNGNYTNGCNGMNGTAYT 802
 Qy 1292 TCTGGCATGCGTGTGAGAGGNTTGTGGAGCGCCCTCTCATACCCAGCACACAC 1351
 Db 803 TYTGGCNGTNCNGARGNGGARITY{WSNTGTGARCCNCCNYTNAH{C{NMGNCAY{CNC 862
 Qy 1352 ACAAGTGTGCTGTTCTGGAGGCGGACCGGCACACTCAAGTGCAGAAAGCCATTGGGGACC 1411
 Db 863 ARMGN{Y{TGGT{Y{NGAR{GNCAR{MNG{CNCNY{TMMGNT{Y{MNG{CNY{TNGGNGAYC 922
 Qy 1412 CCAGCCCTTATCCACTGGTAGCCCGGATGACCCGCTGGTAGGGAACCTCCTCAAGGA 1471
 Db 923 CNGCNCNACNATGAYTGGTNGCNGAY{Y{MNG{Y{TNGTNGNAA{Y{WSN{MNG 982
 Qy 1472 CCGCTGTCTATGACAAATGGCACCTGGACATCTTCA{CACCACATCTCAGGACAGTGGT 1531
 Db 983 CNMNGCNTY{CNCNAA{Y{GNACNY{TNGARA{H{GNG{TNCNCGNCGNCGNAY{GCGGNG 1042
 Qy 1532 CCTTCACTGCAFTGCTGCAATGCTGCGGAGAGGCCACCGCCATGGTGGAGTCTCCA 1591

GenCore version 5.1.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 8, 2003, 15:34:52 ; Search time 175 Seconds
(without alignments)
5509.667 Million cell updates/sec

Title: US-09-831-846-1
Perfect score: 3144
Sequence: 1 gcttgccctctcctcgtga.....gcccgtttgattgattgattg 3144

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues .

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA.*

- 1: /cgn2_6/prodata/2/ina/5A COMB.seq.*
- 2: /cgn2_6/prodata/2/ina/5B COMB.seq.*
- 3: /cgn2_6/prodata/2/ina/6A COMB.seq.*
- 4: /cgn2_6/prodata/2/ina/6B COMB.seq.*
- 5: /cgn2_6/prodata/2/ina/PCTUS COMB.seq.*
- 6: /cgn2_6/prodata/2/ina/backfile1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result NO.	Score	Query Match	Length	ID	Description
1	81.8	2.6	7452	3	US-08-592-500-1
2	81.8	2.6	7452	3	US-08-195-006-1
3	81.8	2.6	7452	5	PCI-US94-07644A-1
4	81.6	2.6	2852	4	US-09-063-950-1
5	79.4	2.5	2019	4	US-09-063-950-3
6	77.2	2.5	5176	4	US-09-182-024A-1
7	63.8	2.0	4843	3	US-08-986-485-1
8	56.4	1.8	2600	2	US-08-427-497E-4
9	56.4	1.8	3189	2	US-08-427-497E-3
10	56.4	1.8	3774	2	US-08-341-843B-1
11	56.4	1.8	3774	2	US-08-427-497E-1
12	56.4	1.8	3774	2	US-08-427-497E-2
13	56.4	1.8	3888	4	US-08-506-296B-13
14	51.6	1.6	4203	2	US-08-866-757-1
15	51.6	1.6	4203	4	US-09-153-593-1
16	50.2	1.6	1685	6	5340934-1
17	48.8	1.6	1362	4	US-08-818-112-7
18	48.8	1.6	1362	4	US-08-818-111-7
19	48.8	1.6	1362	4	US-09-056-556-7
20	48.8	1.6	1362	4	US-09-072-596-7
21	48.8	1.6	4403765	4	US-09-103-840A-2
22	48.8	1.6	4411529	4	US-09-103-840A-1
23	48	1.5	503	4	US-09-297-269-40
24	47.6	1.5	3783	4	US-08-506-296B-20
25	47.4	1.5	906	1	US-07-817-852A-1
26	47.4	1.5	906	1	US-08-470-137-1
27	47.4	1.5	1182	4	US-09-385-028-19

28	47.4	1.5	11604	4	US-09-385-028-13	Sequence 13, Appl
29	47.4	1.5	15079	4	US-09-385-028-1	Sequence 1, Appl
30	47.2	1.5	2338	1	US-08-425-069-1	Sequence 1, Appl
31	47.2	1.5	2338	2	US-08-317-844B-1	Sequence 1, Appl
32	46.6	1.5	44377	2	US-08-804-227C-7	Sequence 7, Appl
33	46.6	1.5	44377	2	US-08-804-198-1	Sequence 1, Appl
34	46.4	1.5	8378	5	PCT-US91-09055-1	Sequence 1, Appl
35	46.2	1.5	4134	4	US-09-162-021B-1	Sequence 1, Appl
36	46.2	1.5	4134	4	US-09-687-477-17	Sequence 17, Appl
37	45.8	1.5	2290	4	US-09-131-648-4	Sequence 4, Appl
38	45.4	1.4	891	4	US-09-452-239-1	Sequence 1, Appl
39	45.4	1.4	4403765	4	US-09-103-840A-2	Sequence 2, Appl
40	45.2	1.4	1018	4	US-09-452-239-35	Sequence 35, Appl
41	44.8	1.4	1119	4	US-08-987-943-2	Sequence 2, Appl
42	44.8	1.4	1755	4	US-08-987-943-1	Sequence 1, Appl
43	44.8	1.4	1926	4	US-09-249-585A-2	Sequence 2, Appl
44	44.8	1.4	2580	3	US-09-050-863-2	Sequence 2, Appl
45	44.8	1.4	2580	4	US-09-359-081-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1
 US-08-592-500-1
 ; Sequence 1, Application US/08592500
 ; Patent No. 6005089
 ; GENERAL INFORMATION:
 ; APPLICANT: Lanza, Francois
 ; APPLICANT: Phillips, David R.
 ; APPLICANT: Cazenave, Jean-Pierre
 ; TITLE OF INVENTION: Platelet Glycoprotein V Gene and Uses
 ; NUMBER OF SEQUENCES: 43
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Townsend and Townsend Kourie and Crew
 ; STREET: 379 Lytton Avenue
 ; CITY: Palo Alto
 ; STATE: California
 ; COUNTRY: US
 ; ZIP: 94301
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/592,500
 ; FILING DATE: 09-JUL-1993
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Dow, Karen B.
 ; REGISTRATION NUMBER: 29,684
 ; REFERENCE/DOCKET NUMBER: 12418-28
 ; TELEPHONE: (415) 326-2400
 ; TELEFAX: (415) 326-2422
 ; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 7452 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: DNA (genomic)
 ; ORIGINAL SOURCE:
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: intron
 ; LOCATION: 1462..2419
 ; NAME/KEY: CDS

LOCATION: 2422..4101
 FEATURE: misc_signal
 NAME/KEY: misc_signal
 LOCATION: 68..76
 OTHER INFORMATION: /function= "Putitive TPA responsive"
 OTHER INFORMATION: element="TRE"
 OTHER INFORMATION: /label= Ets-1
 FEATURE: misc_signal
 NAME/KEY: misc_signal
 LOCATION: 471..478
 OTHER INFORMATION: /function= "Ets-1 cis-acting"
 OTHER INFORMATION: sequence=""
 OTHER INFORMATION: /label= Ets-1
 FEATURE: misc_signal
 NAME/KEY: misc_signal
 LOCATION: 493..502
 OTHER INFORMATION: /function= "Ets-1 cis-acting"
 OTHER INFORMATION: sequence=""
 OTHER INFORMATION: /label= Ets-1
 FEATURE: repeat region
 NAME/KEY: repeat region
 LOCATION: 593..881
 OTHER INFORMATION: /rpt_type= "other"
 OTHER INFORMATION: /label= Alu
 FEATURE: misc_signal
 NAME/KEY: misc_signal
 LOCATION: 897..904
 OTHER INFORMATION: /function= "Ets-1 cis-acting"
 OTHER INFORMATION: sequence=""
 OTHER INFORMATION: /label= Ets-1
 FEATURE: misc binding
 NAME/KEY: misc binding
 LOCATION: 1142..1149
 OTHER INFORMATION: /function= "Spl binding site"
 OTHER INFORMATION: /standard_name= "Spl"
 FEATURE: misc_signal
 NAME/KEY: misc_signal
 LOCATION: 1178..1184
 OTHER INFORMATION: /function= "Ets-1 cis-acting"
 OTHER INFORMATION: sequences=""
 OTHER INFORMATION: /label= Ets-1
 FEATURE: TATA signal
 NAME/KEY: TATA signal
 LOCATION: 1199..1203
 FEATURE: TATA signal
 NAME/KEY: TATA signal
 LOCATION: 1263..1269
 FEATURE: misc binding
 NAME/KEY: misc binding
 LOCATION: 1285..1289
 OTHER INFORMATION: /function= "GATA-1 binding site"
 FEATURE: misc binding
 NAME/KEY: misc binding
 LOCATION: 1321..1326
 OTHER INFORMATION: /function= "GATA-1 binding site"
 FEATURE: misc_signal
 NAME/KEY: misc_signal
 LOCATION: 1365..1372
 OTHER INFORMATION: /function= "Ets-1 cis-acting"
 OTHER INFORMATION: sequences=""
 OTHER INFORMATION: /label= Ets-1
 FEATURE: repeat region
 NAME/KEY: repeat region
 LOCATION: 6133..6440
 OTHER INFORMATION: /rpt_type= "other"
 OTHER INFORMATION: /label= Alu
 FEATURE: misc_signal
 NAME/KEY: misc_signal
 LOCATION: 5610..5615
 OTHER INFORMATION: /standard_name= "Polyadenylation"
 OTHER INFORMATION: signal sequence=""
 FEATURE: misc_signal
 NAME/KEY: misc_signal

LOCATION: 6966..6971
 OTHER INFORMATION: /standard_name= "Polyadenylation"
 OTHER INFORMATION: signal sequence=""
 FEATURE: misc_signal
 NAME/KEY: misc_signal
 LOCATION: 7224..7229
 OTHER INFORMATION: /standard_name= "Polyadenylation"
 OTHER INFORMATION: signal sequence=""
 FEATURE: misc_signal
 NAME/KEY: misc_signal
 LOCATION: 7358..7363
 OTHER INFORMATION: /standard_name= "Polyadenylation"
 OTHER INFORMATION: signal sequence=""
 FEATURE: misc feature
 NAME/KEY: misc feature
 LOCATION: 1..7452
 OTHER INFORMATION: /standard_name= "Nucleotide"
 OTHER INFORMATION: sequence containing the human GPV gene"
 US-08-592-500-1
 Query Match 2.6%; Score 81.8; DB 3; Length 7452;
 Best Local Similarity 50.4%; Pred. No. 1.1e-08;
 Matches 200; Conservative 0; Mismatches 197; Indels 0; Gaps 0;
 QY 707 TGACCCCTGCCAGAAACACCATCAGCCACATCCAGCCCTTTTCCTTTCTGGACCTCGAGA 766
 Db 3233 TGACTCTGTTCCGAGAACCCCGCTGCCAGAGTCCCGGGGTGCTCTTCGGGGAGATGGGG 3292
 QY 767 GCCTCCGCTCCCTGCATCTTGACAGCAATCGGCTGCCAAGCCTTGGGGAGGACACCCCTCC 826
 Db 3293 GCCTGCAGGAGCTGTGGCTGAAACCCAGCTTGGCACCCTGCCCGCCGCGCTTCC 3352
 QY 827 GGGGCTGCTCAACCTGCAGCACCTTATCGTGAACAACAACACAGCTGGGGCGGCTCGCAG 886
 Db 3353 GCACCTGACCGCTGGGTACTTAGGGTGACTGTAGCCCGCGGTGAGCGGCTTC 3412
 QY 887 ATGAGCTTTGAGGACTTCCCTGTGACATTTGGAGGATCTGACCTCTCTTACAACAACC 946
 Db 3413 CGCAGGGCGCTTCCAGGGCTTGGCGAGCTCCAGGTCTCGCCCTGCATCCCAACGGCC 3472
 QY 947 TCCATGGCTGCGGTGGGACTCCGTGCGACGATGTCAACCTCCACAGCTGAGCCTGG 1006
 Db 3473 TGACCGCTTCCCGCAGCGCTTGTGCGGCGCTCGCAAGCTGCCAGGTGCCAGGTGCTCC 3532
 QY 1007 ACCACAACCTGCTGGATCATTCGCCGAGGCGACCTTTGCAGACCTGCAGAACTGGCCC 1066
 Db 3533 GCGGCAACAGGCTGGCGCCCTGCCCGCTTCCGCAATCTCCGCAATCTCAGCAGCTGGAGA 3592
 QY 1067 GCCTGGATCTCACCTCCAAATCGCTGCAGAACTGCC 1103
 Db 3593 GCGTCCAGCTCAGCACCAACACAGCTGGAGACCTGCC 3629
 RESULT 2
 US-08-195-006-1
 ; Sequence 1, Application US/08195006
 ; Patent No. 6083688
 ; GENERAL INFORMATION
 ; APPLICANT: Lanza, Francois
 ; APPLICANT: Phillips, David R.
 ; APPLICANT: Cazenave, Jean-Pierre
 ; TITLE OF INVENTION: Platelet Glycoprotein V Gene and Uses
 ; NUMBER OF SEQUENCES: 43
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Townsend and Townsend Hourie and Crew
 ; STREET: 379 Lytton Avenue
 ; CITY: Palo Alto
 ; STATE: California
 ; COUNTRY: US
 ; ZIP: 94301
 ; COMPUTER READABLE FORM: disk
 ; MEDIUM TYPE: Floppy
 ; COMPUTER: IBM PC compatible

```

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/195,006
FILING DATE: 10-FEB-1994
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/089,455
FILING DATE: 09-JUL-1993
ATTORNEY/AGENT INFORMATION:
NAME: Dow, Karen B.
REGISTRATION NUMBER: 29,684
REFERENCE/DOCKET NUMBER: 12418-28
TELEPHONE: (415) 326-2400
TELEFAX: (415) 326-2422
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 7452 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: intron
LOCATION: 1462..2419
FEATURE:
NAME/KEY: CDS
LOCATION: 2422..4101
FEATURE:
NAME/KEY: misc_signal
LOCATION: 68..76
OTHER INFORMATION: /function= "Putitive TPA responsive"
OTHER INFORMATION: element=
OTHER INFORMATION: /label= TRE
FEATURE:
NAME/KEY: misc_signal
LOCATION: 471..478
OTHER INFORMATION: /function= "Ets-1 cis-acting"
OTHER INFORMATION: sequence=
OTHER INFORMATION: /label= Ets-1
FEATURE:
NAME/KEY: misc_signal
LOCATION: 493..502
OTHER INFORMATION: /function= "Ets-1 cis-acting"
OTHER INFORMATION: sequence=
OTHER INFORMATION: /label= Ets-1
FEATURE:
NAME/KEY: repeat_region
LOCATION: 593..881
OTHER INFORMATION: /rpt_type= "other"
OTHER INFORMATION: /label= Alu
FEATURE:
NAME/KEY: misc_signal
LOCATION: 897..904
OTHER INFORMATION: /function= "Ets-1 cis-acting"
OTHER INFORMATION: sequence=
OTHER INFORMATION: /label= Ets-1
FEATURE:
NAME/KEY: misc_binding
LOCATION: 1142..1149
OTHER INFORMATION: /function= "Sp1 binding site"
OTHER INFORMATION: /standard_name= "Sp1"
FEATURE:
NAME/KEY: misc_signal
LOCATION: 1178..1184
OTHER INFORMATION: /function= "Ets-1 cis-acting"
OTHER INFORMATION: sequences=
OTHER INFORMATION: /label= Ets-1
NAME/KEY: TATA_signal

```

```

LOCATION: 1199..1203
FEATURE:
NAME/KEY: TATA_signal
LOCATION: 1263..1269
FEATURE:
NAME/KEY: misc_binding
LOCATION: 1285..1289
OTHER INFORMATION: /function= "GATA-1 binding site"
FEATURE:
NAME/KEY: misc_binding
LOCATION: 1321..1326
OTHER INFORMATION: /function= "GATA-1 binding site"
FEATURE:
NAME/KEY: misc_signal
LOCATION: 1365..1372
OTHER INFORMATION: /function= "Ets-1 cis-acting"
OTHER INFORMATION: sequences=
OTHER INFORMATION: /label= Ets-1
FEATURE:
NAME/KEY: repeat_region
LOCATION: 6133..6440
OTHER INFORMATION: /rpt_type= "other"
OTHER INFORMATION: /label= Alu
FEATURE:
NAME/KEY: misc_signal
LOCATION: 5610..5615
OTHER INFORMATION: /standard_name= "Polyadenylation"
OTHER INFORMATION: signal sequence=
FEATURE:
NAME/KEY: misc_signal
LOCATION: 6966..6971
OTHER INFORMATION: /standard_name= "Polyadenylation"
OTHER INFORMATION: signal sequence=
FEATURE:
NAME/KEY: misc_signal
LOCATION: 7224..7229
OTHER INFORMATION: /standard_name= "Polyadenylation"
OTHER INFORMATION: signal sequence=
FEATURE:
NAME/KEY: misc_signal
LOCATION: 7358..7363
OTHER INFORMATION: /standard_name= "Polyadenylation"
OTHER INFORMATION: signal sequence=
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1..7452
OTHER INFORMATION: /standard_name= "Nucleotide"
OTHER INFORMATION: sequence containing the human GPV gene"
US-08-195-006-1
Query Match 2.6%; Score 81.8; DB 3; Length 7452;
Best Local Similarity 50.4%; Pred. No. 1.1e-08;
Matches 200; Conservative 0; Mismatches 197; Indels 0; Gaps 0;
QY 707 TGACCCGTTCAGGAACACCATCGCCACATCGCCCTTTTTCCTTTTGGACCTCGAGA 766
Db 3233 TGACTCTGTTTCGAGAACCCCGTGGCAGAGCTCCGGGGGTGCTCTTCGGGGAGATGGGG 3292
QY 767 GCCTCCGCTCCCTGCATTTGACAGCAATCGCTCCCAAGCCTTGGGAGGACACCTCC 826
Db 3293 GCCTCAGGAGCTGTGGTGAACCCACCCAGCTGGCACCCTCCCGCCCGCCCTCC 3352
QY 827 GGGGCTGGTCAACCTCGAGCACCTTATCGTCAACAACAACAGCTGGGCGGATCGCAG 886
Db 3353 GCAACTGAGCCCTCGGTTACTTAGGGGTGACTTGGAGCCCGGCTGAGCCGGCTTC 3412
QY 887 ATGAGCTTTTGGAGCACTTCCTGTGACATTGGAGATCTGGACCTCTCTTCAAACAAC 946
Db 3413 CGCAGGGCCCTTCAGGGCCCTTGGCGAGCTCCAGGTGCTCGCCCTGCACTCAACGGCC 3472
QY 947 TCATGGCCTGCCGTGGGACTCCGTGGACGCATGGTCAACCTCCACCACTGGAGCTGG 1006
Db 3473 TGACCGCCCTCCCGACGGCTTGCTGCGCGGCTCGGCAAGCTGGCCAGGTGTCCCTGC 3532

```

```

OY 1007 ACCCAACCTGCTGATCACATCGCCGAGGGCACCTTTGCAGACTGCAGAAACTGGCCC 1066
DB 3533 GCCGCAACAGGTGGCCGCTGCCCCGCTCCCTTCGGCAATCTCAGCAGCCTGGAGA 3592
OY 1067 GCCTGATCTACCTCCAAATCGCTGGAGAGCTGCC 1103
DB 3593 GCGTCCAGCTGCACCACACACACACACACACACACACACACACACACACACACAC 3629

RESULT 3
PCT-US94-07644A-1
; Sequence 1, Application PC/FUS9407644A
; GENERAL INFORMATION:
; APPLICANT: COR Therapeutics, Inc.
; TITLE OF INVENTION: Platelet Glycoprotein V Gene and Uses
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Hourie and Crew
; STREET: 379 Lytton Avenue
; CITY: Palo Alto
; STATE: California
; COUNTRY: US
; ZIP: 94301
; COMPUTER READABLE FORM: disk
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/07644A
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Dow, Karen B.
; REGISTRATION NUMBER: 29,684
; REFERENCE/DOCKET NUMBER: 012418-003000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 326-2400
; TELEFAX: (415) 326-2422
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7452 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: intron
; LOCATION: 1462..2419
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 2422..4101
; FEATURE:
; NAME/KEY: misc signal
; LOCATION: 68..76
; OTHER INFORMATION: /function= "Putitive TPA responsive
; OTHER INFORMATION: element"
; OTHER INFORMATION: /label= TRE
; FEATURE:
; NAME/KEY: misc signal
; LOCATION: 471..478
; OTHER INFORMATION: /function= "Ets-1 cis-acting
; OTHER INFORMATION: sequence"
; OTHER INFORMATION: /label= Ets-1
; FEATURE:
; NAME/KEY: misc signal
; LOCATION: 493..502
; OTHER INFORMATION: /function= "Ets-1 cis-acting
; OTHER INFORMATION: sequence"
; OTHER INFORMATION: /label= Ets-1
; FEATURE:
; NAME/KEY: misc signal
; LOCATION: 1..7452
; OTHER INFORMATION: /standard_name= "Nucleotide
; OTHER INFORMATION: sequence Containing the human Gpv gene"
; OTHER INFORMATION:
; PCT-US94-07644A-1

```

```

; NAME/KEY: repeat region
; LOCATION: 593..881
; OTHER INFORMATION: /rpt_type= "other"
; OTHER INFORMATION: /label= Alu
; FEATURE:
; NAME/KEY: misc signal
; LOCATION: 897..904
; OTHER INFORMATION: /function= "Ets-1 cis-acting
; OTHER INFORMATION: sequence"
; OTHER INFORMATION: /label= Ets-1
; FEATURE:
; NAME/KEY: misc binding
; LOCATION: 1142..1149
; OTHER INFORMATION: /function= "Spi binding site."
; OTHER INFORMATION: /standard_name= "Spi"
; FEATURE:
; NAME/KEY: misc signal
; LOCATION: 1178..1184
; OTHER INFORMATION: /function= "Ets-1 cis-acting
; OTHER INFORMATION: sequences"
; OTHER INFORMATION: /label= Ets-1
; FEATURE:
; NAME/KEY: TATA signal
; LOCATION: 1199..1203
; FEATURE:
; NAME/KEY: TATA signal
; LOCATION: 1263..1269
; FEATURE:
; NAME/KEY: misc binding
; LOCATION: 1285..1289
; OTHER INFORMATION: /function= "GATA-1 binding site"
; OTHER INFORMATION:
; NAME/KEY: misc binding
; LOCATION: 1321..1326
; OTHER INFORMATION: /function= "GATA-1 binding site"
; FEATURE:
; NAME/KEY: misc signal
; LOCATION: 1365..1372
; OTHER INFORMATION: /function= "Ets-1 cis-acting
; OTHER INFORMATION: sequences"
; OTHER INFORMATION: /label= Ets-1
; FEATURE:
; NAME/KEY: repeat region
; LOCATION: 6133..6440
; OTHER INFORMATION: /rpt_type= "other"
; OTHER INFORMATION: /label= Alu
; FEATURE:
; NAME/KEY: misc signal
; LOCATION: 5610..5615
; OTHER INFORMATION: /standard_name= "Polyadenylation
; OTHER INFORMATION: signal sequence"
; FEATURE:
; NAME/KEY: misc signal
; LOCATION: 6966..6971
; OTHER INFORMATION: /standard_name= "Polyadenylation
; OTHER INFORMATION: signal sequence"
; FEATURE:
; NAME/KEY: misc signal
; LOCATION: 7224..7229
; OTHER INFORMATION: /standard_name= "Polyadenylation
; OTHER INFORMATION: signal sequence"
; FEATURE:
; NAME/KEY: misc signal
; LOCATION: 7358..7363
; OTHER INFORMATION: /standard_name= "Polyadenylation
; OTHER INFORMATION: signal sequence"
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1..7452
; OTHER INFORMATION: /standard_name= "Nucleotide
; OTHER INFORMATION: sequence Containing the human Gpv gene"
; OTHER INFORMATION:
; PCT-US94-07644A-1

```

Query Match 2.6%; Score 81.8; DB 5; Length 7452;
 Best Local Similarity 50.4%; Pred. No. 1.1e-08;
 Matches 200; Conservative 0; Mismatches 197; Indels 0; Gaps 0;

QY 707 TCACCTGTCCAGGAACACCATCAGCCACATCAGCCCTTTTCCCTTTTGGACCTCGAGA 766
 |||||
 Db 3233 TGACTCTGTTGAGAAACCGCTGGCAGAGCTCCGGGGGTGCTCTTCGGGGAGATGGGG 3292

QY 767 GCCTCCGCTCCCTGTCATTTGACGCAATCGCTGCCAAGCCTTGGGGAGGACACCCCTCC 826
 |||||
 Db 3293 GCCTGAGGAGCTGTGGCTGAACCGCACCCAGCTGGCACCTTGCCTGGCCGCGCTTCC 3352

QY 827 GGGGCTGTCAACCTGAGACCTTATCTGTGAACAAACACAGCTGGCCGCGCATCGCAG 886
 |||||
 Db 3353 GCAACCTGAGCCGCTTGGGGTACTTAGGGTACTCTGAGCCGCGGCTGAGCGGCTTC 3412

QY 887 ATGAGCTTTTTCAGACTTTCCTGCTGACATTTGAGATCTGACCTCTCTACAACAACC 946
 |||||
 Db 3413 CCAGGGCCCTTCCAGGGCCTTGGGAGCTCCAGGTGCTGCCCTGCACTCCAAGGCC 3472

QY 947 TCCATGGCTGCGGTGGACTCCGTCGACGATGTCAACTCCACACAGCTGAGCCTGG 1006
 |||||
 Db 3473 TGACCCGCTCCCGCAGGCTTGTCTGGGGCTCGGCAAGCTGGCCAGGTTCCTTGC 3532

QY 1007 ACCACAACTGTGATACATCGCCGAGGACCTTTGACACCTGCAAACTGGCCCC 1066
 |||||
 Db 3533 GCCGCAACAGGCTGGCGCCCTGCCCCGTGCCCTTTCCGCAATCTCAGCAGCCTGGAGA 3592

QY 1067 GCTGTGATCTACCTCCATCGCTGACAGCTGCC 1103
 |||||
 Db 3593 GGGTCCAGCTCAGCCAAACAGCTGGAGACCTGCC 3629

RESULT 4

US-09-063-950-1
 ; Sequence 1, Application US/09063950C
 ; Patent No. 6225085
 ; GENERAL INFORMATION:
 ; APPLICANT: Holtzman, Douglas A.
 ; TITLE OF INVENTION: NOVEL LRSG PROTEIN AND NUCLEIC ACID MOLECULES AND USES
 ; FILE REFERENCE: MEI-019
 ; CURRENT APPLICATION NUMBER: US/09/063,950C
 ; CURRENT FILING DATE: 1998-04-21
 ; NUMBER OF SEQ ID NOS: 9
 ; SOFTWARE: Patent In Ver. 2.0
 ; SEQ ID NO 1
 ; LENGTH: 2852
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (160)..(2178)
 ; US-09-063-950-1

Query Match 2.6%; Score 81.6; DB 4; Length 2852;
 Best Local Similarity 46.6%; Pred. No. 8.9e-09;
 Matches 338; Conservative 0; Mismatches 379; Indels 9; Gaps 2;

QY 315 CCGAGCCCGGGAGGTTAGCTCGAGTCTGTCTCGGGGGGAAAGATGGCTGGCCGA 374
 |||||
 Db 15 CCGGAGCCCGGGGGTGGACGCGGACTCGAACGAGTTGTTCCGGGACCCAGGACCCC 74

QY 375 GCGGGGAGCCCGGGGCCCCCGGCGCTCGGTGCCACCCAGCCGCGGGGTAGATG 434
 |||||
 Db 75 CTGGGCCCGCAGCCCGCAGAAAGACTGAGGCGCGGCGCTGCCCGCCGCTCCCTGGC 134

QY 435 CTGCTCCGCCAGGCGCTGAGTGACACGACCATGAGACCTGCTTGGTGGCTGTGCTAGC 494
 |||||
 Db 135 CCGCCGCGCTCCCG- ----GGACAGAAGATGCTCCAGGTCCCTCTGCTGCTGCC 188

QY 495 GTTTGGCATGGGCTTTGGCGTGGTCGACCGCTGCCCCAAAGTACTGTCTGCCAGAACT 554
 |||||

Db 189 GCTGCTCTGCTACTGGCCCTTGGGCTGGGGTGCAGGGCTGCCCATCCGGCTGCCAGTG 248

QY 555 GTCTGAGTCACTGGGGACCTGTGCCCTCCAAGGGGTGCTCTTTGTACCCCTGATAT 614
 |||||

Db 249 CAGCCAGCCACAGACAGTCTTCTGACCTGCCCGCCAGGGACCCAGGTGCCCGAGACT 308
 |||||

QY 615 TGACCCGGGAGACAGTGGAGCTGGCTGGGGGCACTTTCATCATCCACATCAGCGGCCA 674
 |||||

Db 309 GCCACCCGACACCGTGGGGTGTACGTTCTTTGAGAAACGGCATCACCATGCTCGACGCG 368
 |||||

QY 675 GGACTTTCCAAACATGACGGGCTGGTGGACCTGACCTGACCTGTCAGGAACACCATCAGCCA 734
 |||||

Db 369 CAGCTTTCCGGCTCCGGGCTGCAGCTCTGACCTGTGACACCTGTGACAGAACAGATCGCCAG 428
 |||||

QY 735 CATCAGGCCCTTTTCTTTTGGACCTCGAGACCTCCGCTCCCTGCACTTTTGACAGCAA 794
 |||||

Db 429 CCTGCCAGCGGGTCTTCCAGCCACTCGCAACCTCAGCAACCTGGACCTGACGGCCAA 488
 |||||

QY 795 TGGCTGCCAAGCCTTGGGGAGGACCCCTCGGGGCTGTGTAACCTGACCTGACGACCTTAT 854
 |||||

Db 489 CAGGCTGATGAAATACCAATGAGACCTTCCGTGGCTGGGGCTCGAGCGCTCTA 548
 |||||

QY 855 CGTGAACAAACACAGCTGGGGGATCGAGATGAGCTTTTGAAGACTTCTCTGCTGAC 914
 |||||

Db 549 CCTGGCAAGAACCCGATCCGCCACATCCAGCTGTGCTTCGACAGCTCGACCGCT 608
 |||||

QY 915 ATTGAGGATCTGGACCTCTCTCAACAACCTCCATGGCTGCGCTGGGACTCCGTTGG 974
 |||||

Db 609 CTGGAG--CTCAGCTGCAGGACACAGAGCTGGGGCACTGCCCGCTGGCGCTGCC 665
 |||||

QY 975 AGCATGTCAAACCTCCACAGCTGAGCTGGACCAACACCTGCTGGATCAATCGCCGA 1034
 |||||

Db 666 CCGCTGTGCTGGACTCGACCAACAACAGCTTCTGGCCCTGGAGCCCGGATCCT 725

QY 1035 GGGCAC 1040
 |||||

Db 726 GGACAC 731
 |||||

RESULT 5

US-09-063-950-3
 ; Sequence 3, Application US/09063950C
 ; Patent No. 6225085
 ; GENERAL INFORMATION:
 ; APPLICANT: Holtzman, Douglas A.
 ; TITLE OF INVENTION: NOVEL LRSG PROTEIN AND NUCLEIC ACID MOLECULES AND USES
 ; FILE REFERENCE: MEI-019
 ; CURRENT APPLICATION NUMBER: US/09/063,950C
 ; CURRENT FILING DATE: 1998-04-21
 ; NUMBER OF SEQ ID NOS: 9
 ; SOFTWARE: Patent In Ver. 2.0
 ; SEQ ID NO 3
 ; LENGTH: 2019
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (1)..(2019)
 ; US-09-063-950-3

Query Match 2.5%; Score 79.4; DB 4; Length 2019;
 Best Local Similarity 49.3%; Pred. No. 2.4e-08;
 Matches 237; Conservative 0; Mismatches 241; Indels 3; Gaps 1;

QY 560 AGTCACTGGGACCTGTGCCCTCCAAAGGGGCTCTTTGTACCCCTGATTTGACC 619
 |||||

Db 95 AGCCACAGACAGTCTTCTGCACTGCCCGCAGGGGACCCAGGTGCCCGAGAGCTGCCAC 154
 |||||

QY 620 GCGGACAGTGGAGCTGGCCCTGGGGCACTTTCATCATCCACATCAGCCCGCAGGACT 679
 |||||

Db 155 CCGACACCGTGGGGTGTACGTTCTTTGAGAACGGCATCACCATGCTCGACGCGAGCAGCT 214
 |||||

QY 804 AAGCCCTGGGGAGCACCCCTCGGGCCCTGGTCAACCTGCGACACCTTTATCGTGAACAA 863
 Db 789 CAAACTGACAGATGGGGCCCTTCTGGGCACTGTCCAAAATGCAATGCTGCTCACCTGGAGTA 848
 QY 864 CAACAGCTGGGGGATCGCAGATGAGGCTTTTGGAGACTTCTGCTGACATTTGGAGGA 923
 Db 849 GCACAGCTGGTAGAAGTGAACAGCGGCTCGCTTACGGCCTCACGG---CCCTGCATCA 905
 QY 924 TCTGGACCTCTCTCAACAACCTCCATGGCTGGCTGGGACTCGCTGGAGCGCATGCT 983
 Db 906 GCTCCACCTCAGCAACAATTCATCGCTCGCATTCACCCGAAGGGCTGAGGCTTTCGCCA 965
 QY 984 CAACCTCCACAGCTGAGCTGGACCAACCTGCTGGATCAATCGCCGAGGGCACCTT 1043
 Db 966 GAAGCTGATGATGGTCTGCTTCAACAACCTGACACCGCTGGAGGAGAGAGCCT 1025
 QY 1044 TCGACACTCGAAGAACTGGCCCGCTCGATCTCACCTCAAT 1086
 Db 1026 GCCCGAGCTGAGCAGCCTGAGTGTCTGCTGCTCAGCCCAAT 1068

RESULT 8

US-08-427-497E-4
 ; Sequence 4, Application US/08427497E
 ; Patent No. 5969124
 ; GENERAL INFORMATION:
 ; APPLICANT: Lemmon, Vance
 ; TITLE OF INVENTION: A Method for Characterizing the
 ; TITLE OF INVENTION: Nucleotide Sequence of L1CAM and
 ; Patent No. 5969124
 ; TITLE OF INVENTION: the Nucleotide Sequence
 ; TITLE OF INVENTION: Characterized Thereby
 ; NUMBER OF SEQUENCES: 44
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Fay, Sharpe, Beall, Pagan,
 ; ADDRESSEE: Minnich & Mckee
 ; STREET: 1100 Superior Avenue
 ; STREET: Suite 700
 ; CITY: Cleveland
 ; STATE: Ohio
 ; COUNTRY: U.S.A.
 ; ZIP: 44114-2518

; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette, 3.50 inch, 720 Kb
 ; MEDIUM TYPE: storable
 ; COMPUTER: Compaq Prolinea 5100e
 ; OPERATING SYSTEM: DOS 5.0
 ; SOFTWARE: ASCII
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/427,497E
 ; FILING DATE: April 24, 1995
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 07/904,991
 ; FILING DATE: June 26, 1992
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Minnich, Richard J.
 ; REGISTRATION NUMBER: 24,175
 ; REFERENCE/DOCKET NUMBER: CWR 2 149-3-1
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (216) 861-5582
 ; TELEFAX: (216) 241-1666
 ; TELEX: (216) 980162
 ; INFORMATION FOR SEQ ID NO: 4:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 2600
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: nucleic acids
 ; HYPOTHETICAL: irrelevant
 ; ANTI-SENSE: no

ORIGINAL SOURCE:
 ORGANISM: homo sapiens
 INDIVIDUAL ISOLATE: 17-18 week fetus
 IMMEDIATE SOURCE:
 LIBRARY: Stratagene cDNA Library 936206
 CLONE: 4
 PUBLICATION INFORMATION:
 AUTHORS: Hlavin, Mary Louise
 AUTHORS: Lemmon, Vance
 TITLE: Molecular structure and functional testing of
 TITLE: human L1CAM: an interspecies comparison.
 JOURNAL: GENOMICS
 VOLUME: 11
 ISSUE:
 PAGES: 416-423
 DATE: 1991
 RELEVANT RESIDUES IN SEQ ID NO: 1108 to 3708
 US-08-427-497E-4

Query Match 1.8%; Score 56.4; DB 2; Length 2600;
 Best Local Similarity 49.3%; Pred. No. 0.0026;
 Matches 147; Conservative 0; Mismatches 151; Indels 0; Gaps 0;
 QY 1334 TCATCCCAGCACACAAAGTTGCTGGTTCAGGGCCAGGGCCACACTCAAGT 1393
 Db 176 TCCTGACTCGGCAATCAGACGTACATGGCTGTCCAGGGCAGCACTGCCTTCTGT 235
 QY 1394 GCAAAGCCATTGGGGACCCGCCCTTATCCACTGGGTAGCCCGGATGACCCGCTGG 1453
 Db 236 GCAAGGCCCTTCGGAGGGCTGTGCCAGTGTTTCAGTGGCTGGAGGATGGGCAACAG 295
 QY 1454 TAGGAACTCTCAAGAGCCGCTGTATGACAAATGGCACCTGGACATCTTTCATCACCA 1513
 Db 296 TGCTTCAGGACCAACGCTTCTCCCTATGCCAATGGGACCTGGGCAATTCGAGACTCC 355
 QY 1514 CATCTCAGCAGAGTGGTCCCTTACCTGCAATGCTGCCAATGCTCCGGAGAGCCACCG 1573
 Db 356 AGGCCAATGACACCGGACGCTACTTCTGCTGGTCCCAATGACCAAAAACAATGTTACCA 415
 QY 1574 CCATGTTGGAGTCTCCATCGCTCCAGCTGCCACACTCAGCAACAGCACCCGCGCAC 1631
 Db 416 TCATGGCTAACCTGAAGGTTAAAGATGCAACTCAGATCACTCAGGGGCCCCGCGCAC 473

RESULT 9

US-08-427-497E-3
 ; Sequence 3, Application US/08427497E
 ; Patent No. 5969124
 ; GENERAL INFORMATION:
 ; APPLICANT: Lemmon, Vance
 ; TITLE OF INVENTION: A Method for Characterizing the
 ; TITLE OF INVENTION: Nucleotide Sequence of L1CAM and
 ; Patent No. 5969124
 ; TITLE OF INVENTION: the Nucleotide Sequence
 ; TITLE OF INVENTION: Characterized Thereby
 ; NUMBER OF SEQUENCES: 44
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Fay, Sharpe, Beall, Pagan,
 ; ADDRESSEE: Minnich & Mckee
 ; STREET: 1100 Superior Avenue
 ; STREET: Suite 700
 ; CITY: Cleveland
 ; STATE: Ohio
 ; COUNTRY: U.S.A.
 ; ZIP: 44114-2518
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette, 3.50 inch, 720 Kb
 ; MEDIUM TYPE: storable
 ; COMPUTER: Compaq Prolinea 5100e
 ; OPERATING SYSTEM: DOS 5.0
 ; SOFTWARE: ASCII
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/427,497E

```

TITLE OF INVENTION: Nucleotide Sequence of L1CAM and
Patent No. 5872225
TITLE OF INVENTION: the Nucleotide Sequence
TITLE OF INVENTION: Characterized Thereby
NUMBER OF SEQUENCES: 39
CORRESPONDENCE ADDRESS:
ADDRESS: Fay, Sharpe, Beall, Fagan,
ADDRESS: Minnich & McKee
STREET: 1100 Superior Avenue
STREET: Suite 700
CITY: Cleveland
STATE: Ohio
COUNTRY: U.S.A.
ZIP: 44114-2518
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 720 Kb
MEDIUM TYPE: storable
COMPUTER: Compaq Prolinea 5100e
OPERATING SYSTEM: DOS 5.0
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/341,843B
FILING DATE: No. 5872225ember 18, 1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/904,991
FILING DATE: June 26, 1992
ATTORNEY/AGENT INFORMATION:
NAME: Minnich, Richard J.
REGISTRATION NUMBER: 24,175
REFERENCE/DOCKET NUMBER: CWR 2 149-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (216) 861-5582
TELEFAX: (216) 241-1666
TELEX: (216) 980162
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 3774
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: irrelevant
ANTI-SENSE: no
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
INDIVIDUAL ISOLATE: 17-18 week fetus
IMMEDIATE SOURCE:
LIBRARY: Stratagene cDNA Library 936206
CLONE: synthesis of 4 clones
PUBLICATION INFORMATION:
AUTHORS: Hlavir, Mary Louise
AUTHORS: Lemmon, Vance
TITLE: Molecular structure and functional
TITLE: testing of human L1CAM: an
TITLE: interspecies comparison.
JOURNAL: GENOMICS
VOLUME: 11
ISSUE: 11
PAGES: 416-423
DATE: 1991
RELEVANT RESIDUES IN SEQ ID NO: 1 to 3774
US-08-341-843B-1
Query Match 1.8%; Score 56.4; DB 2; Length 3774;
Best Local Similarity 49.3%; Pred. No. 0.0028;
Matches 147; Conservative 0; Mismatches 151; Indels 0; Gaps 0;
Qy 1334 TCATCACCAGCACACACAAAGTTCTGGAGGGCCAGGGCCACACTCAAGT 1393
Db 1283 TCCTGACTGGGACAAATCAGACCTACATGGCTCTCCAGGGCAGCAGCTGCTTCTGT 1342
Qy 1394 GCAAAGCCATTGGGGACCCCGCCCTTATCCACTGGGTAGCCCGCCGATGACCGCTGG 1453

```

```

FILING DATE: April 24, 1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/904,991
FILING DATE: June 26, 1992
ATTORNEY/AGENT INFORMATION:
NAME: Minnich, Richard J.
REGISTRATION NUMBER: 24,175
REFERENCE/DOCKET NUMBER: CWR 2 149-3-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (216) 861-5582
TELEFAX: (216) 241-1666
TELEX: (216) 980162
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 3189
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: nucleic acids
HYPOTHETICAL: irrelevant
ANTI-SENSE: no
ORIGINAL SOURCE:
ORGANISM: homo sapiens
INDIVIDUAL ISOLATE: 17-18 week fetus
IMMEDIATE SOURCE:
LIBRARY: Stratagene cDNA Library 936206
CLONE: 3.1
PUBLICATION INFORMATION:
AUTHORS: Hlavir, Mary Louise
AUTHORS: Lemmon, Vance
TITLE: Molecular structure and functional testing of
TITLE: human L1CAM: an interspecies comparison.
JOURNAL: GENOMICS
VOLUME: 11
ISSUE: 11
PAGES: 416-423
DATE: 1991
RELEVANT RESIDUES IN SEQ ID NO: 548 to 3736
US-08-427-497E-3
Query Match 1.8%; Score 56.4; DB 2; Length 3189;
Best Local Similarity 49.3%; Pred. No. 0.0028;
Matches 147; Conservative 0; Mismatches 151; Indels 0; Gaps 0;
Qy 1334 TCATCACCAGCACACACAAAGTTCTGGAGGGCCAGGGCCACACTCAAGT 1393
Db 736 TCCTGACTGGGACAAATCAGACCTACATGGCTCTCCAGGGCAGCAGCTGCTTCTGT 795
Qy 1394 GCAAAGCCATTGGGGACCCCGCCCTTATCCACTGGGTAGCCCGCCGATGACCGCTGG 1453
Db 796 GCAAGGCCCTTGGAGCCCTGTGCCAGTGTTCAGTGGCTGGACGAGGATGGGACACAG 855
Qy 1454 TAGGGAATCTCTCAAGSACCGGTGCTATGACAAATGGACCCCTGGACATTTTCACCA 1513
Db 856 TGCTTCAGACGACGACGCTTCTTCCCTATGCCAATGGGACCCCTGGCAATTCAGACCTCC 915
Qy 1514 CATCTCAGGACAGTGGTCCCTTACCTGATGTCCTGCAATGTCCTGCGGAGAGCCAGG 1573
Db 916 AGGCCAATGACACCCGAGCCGCTACTTCTGCTGGCTGCTGCAATGACCAAAACAAATGTACCA 975
Qy 1574 CCATGCTGGAGGTCCTCCATGCTCCAGCTGCCACACTCAGCAACAGCACACCGCCGAC 1631
Db 976 TCATGGCTAACCTGAAGGTTAAGATGCAACTCAGATCACTCAGGGGCCCGCCAGCAC 1033
RESULT 10
US-08-341-843B-1
; Sequence 1, Application US/08341843B
; Patent No. 5872225
; GENERAL INFORMATION:
; APPLICANT: Lemmon, Vance
; TITLE OF INVENTION: A Method for Characterizing the

```


Db 1343 GCAAGGCCTTCGGAGCGCTGTGCCAGTGTTTCAGTGGCTGGACGAGGATGGGACAACAG 1402
 Qy 1454 TAGGGAACCTCCAAGGACCGCTGTATGACAAATGGCACCTTGACATCTTTCATCACCA 1513
 Db 1403 TCCCTTCAGGACGAAGCTTCTCCCTATGACAAATGGGACCTTGGAATTCAGACCTCC 1462
 Qy 1514 CATCTCAGGACAGTGGTCCCTTACCTGCAATGCTGCCAATGCTCCGGAGAGCCACGG 1573
 Db 1463 AGGCCAATGACACCGGAGCTACTTCTGCCTGGCTGACCAATGACCAAAAATGTTACCA 1522
 Qy 1574 CCATGCTGGAGTCTCCATCTGTCAGCTGCCACACTCAGCAACAGCAGCACCACGGC 1631
 Db 1523 TCATGGCTAACCTGAAGTAAAGTCAACTCAGATCACTCAGGGGGCCCCCGCAGCAC 1580

RESULT 11
 US-08-427-497E-1
 ; Sequence 1, Application US/08427497E
 ; Patent No. 5969124
 ; GENERAL INFORMATION:
 ; APPLICANT: Lemmon, Vance
 ; TITLE OF INVENTION: A Method for Characterizing the
 ; Nucleotide Sequence of L1CAM and
 ; Patent No. 5969124
 ; TITLE OF INVENTION: the Nucleotide Sequence
 ; Characterized Thereby
 ; NUMBER OF SEQUENCES: 44
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Fay, Sharpe, Beall, Fagan,
 ; STREET: 1100 Superior Avenue
 ; CITY: Cleveland
 ; STATE: Ohio
 ; COUNTRY: U.S.A.
 ; ZIP: 44114-2518
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette, 3.50 inch, 720 Kb
 ; MEDIUM TYPE: .storable
 ; COMPUTER: Compaq Prolinea 5100e
 ; OPERATING SYSTEM: DOS 5.0
 ; SOFTWARE: ASCII
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/427,497E
 ; FILING DATE: June 26, 1992
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Minnich, Richard J.
 ; REGISTRATION NUMBER: 24,175
 ; REFERENCE/DOCKET NUMBER: CWR 2 149-3-1
 ; TELEPHONE: (216) 861-5582
 ; TELEFAX: (216) 241-1666
 ; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 3774
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: nucleic acids
 ; HYPOTHETICAL: irrelevant
 ; ANTI-SENSE: no
 ; ORGANISM: Homo Sapiens
 ; INDIVIDUAL ISOLATE: 17-18 week fetus
 ; IMMEDIATE SOURCE:
 ; LIBRARY: Stratagene cDNA Library 936206
 ; CLONE: synthesis of 4 clones

PUBLICATION INFORMATION:
 AUTHORS: Hlavin, Mary Louise
 AUTHORS: Lemmon, Vance
 TITLE: Molecular structure and functional
 TITLE: testing of human L1CAM: an
 TITLE: interspecies comparison.
 JOURNAL: GENOMICS
 VOLUME: 11
 ISSUE:
 PAGES: 416-423
 DATE: 1991
 RELEVANT RESIDUES IN SEQ ID NO: 1 to 3774
 US-08-427-497E-1
 Query Match 1.8%; Score 56.4; DB 2; Length 3774;
 Best Local Similarity 49.3%; Pred. No. 0.0029;
 Matches 147; Conservative 0; Mismatches 151; Indels 0; Gaps 0;
 Qy 1334 TCATCACCCAGCACACACAAAGTTGCTTCTGGAGGGCCAGGGCCACACTCAAGT 1393
 Db 1283 TCCCTGACTGGGACAAATCAGACGTAATGGCTGTCCAGGGCAGCACTGCCCTTCTGT 1342
 Qy 1394 GCAAGCCATTTGGGGACCCACCCCTTATCCACTGGTATGCCCTGGTACCCCGGACCGCTGG 1453
 Db 1343 GCAGGCCCTTCGGAGCGCTGTGCCAGTTCAGTGTGCTGGACGAGGATGGGACAACAG 1402
 Qy 1454 TAGGGAACCTCCTCAAGGACCGCTGTCTATGACAAATGGCACCTTTCATCACCA 1513
 Db 1403 TGCCTCAGGACGAAGCTTCTCCCTATGCCAATGGGACCTTGGCATTTCGAGACCTCC 1462
 Qy 1514 CATCTCAGGACAGTGGTCCCTTACCTGCAATGCTGCCAATGCTCCGGAGAGCCACGG 1573
 Db 1463 AGGCCAATGACACCGGAGCTACTTCTGCCTGGCTGCCAATGACCAAAAATGTTACCA 1522
 Qy 1574 CCATGCTGGAGTCTCCATCTGTCAGCTGCCACACTCAGCAACAGCAGCACCACGGC 1631
 Db 1523 TCATGGCTAACCTGAAGTAAAGTCAACTCAGATCACTCAGGGGGCCCCCGCAGCAC 1580
 RESULT 12
 US-08-427-497E-2
 ; Sequence 2, Application US/08427497E
 ; Patent No. 5969124
 ; GENERAL INFORMATION:
 ; APPLICANT: Lemmon, Vance
 ; TITLE OF INVENTION: A Method for Characterizing the
 ; Nucleotide Sequence of L1CAM and
 ; Patent No. 5969124
 ; TITLE OF INVENTION: the Nucleotide Sequence
 ; Characterized Thereby
 ; NUMBER OF SEQUENCES: 44
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Fay, Sharpe, Beall, Fagan,
 ; ADDRESS: Minnich & McKee
 ; STREET: 1100 Superior Avenue
 ; CITY: Cleveland
 ; STATE: Ohio
 ; COUNTRY: U.S.A.
 ; ZIP: 44114-2518
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette, 3.50 inch, 720 Kb
 ; MEDIUM TYPE: .storable
 ; COMPUTER: Compaq Prolinea 5100e
 ; OPERATING SYSTEM: DOS 5.0
 ; SOFTWARE: ASCII
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/427,497E
 ; FILING DATE: April 24, 1995
 ; CLASSIFICATION: 435
 ; PRIORITY DATA:
 ; APPLICATION NUMBER: 07/904,991
 ; FILING DATE: June 26, 1992

ADDRESS: The Scripps Research Institute
 STREET: 10550 No. 6313265th Torrey Pines Road, TPC-8
 CITY: La Jolla
 STATE: California
 COUNTRY: U.S.
 ZIP: 92037
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent Release #1.0, Version #1.25
 CURRENT APPLICATION NUMBER: US/08/506,296B
 FILING DATE: 24-JUL-1995
 CLASSIFICATION: 514
 ATTORNEY/AGENT INFORMATION:
 NAME: Fitting, Thomas
 REGISTRATION NUMBER: 34,163
 REFERENCE/DOCKET NUMBER: TSRI 488.0
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (619) 554-2937
 TELEFAX: (619) 554-6312
 INFORMATION FOR SEQ ID NO: 13:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 3888 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 NAME/KEY: CDS
 LOCATION: 12..3773
 US-08-506-296B-13

ATTORNEY/AGENT INFORMATION:
 NAME: Minnich, Richard J.
 REGISTRATION NUMBER: 24,175
 REFERENCE/DOCKET NUMBER: CWR 2 149-3-1
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (216) 861-5582
 TELEFAX: (216) 241-1666
 TELEX: (216) 980162
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 3774
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: nucleic acids
 HYPOTHETICAL: irrelevant
 ANTI-SENSE: no
 ORIGINAL SOURCE:
 ORGANISM: Homo Sapiens
 INDIVIDUAL ISOLATE: 17-18 week fetus
 IMMEDIATE SOURCE:
 LIBRARY: Stratagene cDNA Library 936206
 CLONE: synthesis of 4 clones
 PUBLICATION INFORMATION:
 AUTHORS: Hlavin, Mary Louise
 TITLE: Molecular structure and functional testing of human L1CAM: an interspecies comparison.
 JOURNAL: GENOMICS
 VOLUME: 11
 ISSUE:
 PAGES: 416-423
 DATE: 1991
 RELEVANT RESIDUES IN SEQ ID NO: 1 to 3774.
 US-08-427-497E-2

Query Match 1.8%; Score 56.4; DB 4; Length 3888;
 Best Local Similarity 49.3%; Pred. No. 0.0029;
 Matches 147; Conservative 0; Mismatches 151; Indels 0; Gaps 0;

QY 1334 TCATCACCAGCACACACAAGTTGCTGTTGGAGGGCCAGCGCCACACTCAAGT 1393
 Db 1294 TCCTGACTGGGACAAATCAGACTACATGGCTGCTCCAGGGAGCAGCTGCCTTCTGT 1353
 QY 1394 GCAAAAGCCATTGGGACCCAGCCCTTATCCACTGGGTAGCCCCCGATGACCGCCCTGG 1453
 Db 1354 GCAAGGCCCTTCGGAGCGCTGTGCCAGTGTTCAGTGTGCTGACGAGATGGACACAG 1413
 QY 1454 TAGGGAACCTCCTCAAGGACCGCTGTCTATGACAATGGCACCCCTGGACATCTTCATCACA 1513
 Db 1414 TCGTTCCAGGACGAAACGCTTCTCCCTATGCCAATGGGACCCCTGGCATTCGAGACCTCC 1473
 QY 1514 CATCTCAGGACAGTGTGCTTCCACTGCAATGTGTCGCAATGCTGCCGAGAGCCACGG 1573
 Db 1474 AGGCCAATGACACCGGACGCTACTTCTGCTGGTCCCAATGACCAAAACAATGTTACCA 1533
 QY 1574 CCATGTGAGGTTCCATCGTCCAGCTGCCACACCTCAGCAACAGCACCGCCGAC 1631
 Db 1534 TCATGGCTAACCTGAAGGTTAAAGATGCAACTCAGATCACTCAGGGGCCCCCGACAGC 1591

Query Match 1.8%; Score 56.4; DB 2; Length 3774;
 Best Local Similarity 49.3%; Pred. No. 0.0029;
 Matches 147; Conservative 0; Mismatches 151; Indels 0; Gaps 0;

QY 1334 TCATCACCAGCACACACAAGTTGCTGTTGGAGGGCCAGCGCCACACTCAAGT 1393
 Db 1283 TCCTGACTGGGACAAATCAGACTACATGGCTGCTCCAGGGAGCAGCTGCCTTCTGT 1342
 QY 1394 GCAAAAGCCATTGGGACCCAGCCCTTATCCACTGGGTAGCCCCCGATGACCGCCCTGG 1453
 Db 1343 GCAAGGCCCTTCGGAGCGCTGTGCCAGTGTTCAGTGTGCTGACGAGATGGACACAG 1402
 QY 1454 TAGGGAACCTCCTCAAGGACCGCTGTCTATGACAATGGCACCCCTGGACATCTTCATCACA 1513
 Db 1403 TCGTTCCAGGACGAAACGCTTCTCCCTATGCCAATGGGACCCCTGGCATTCGAGACCTCC 1462
 QY 1514 CATCTCAGGACAGTGTGCTTCCACTGCAATGTGTCGCAATGCTGCCGAGAGCCACGG 1573
 Db 1463 AGGCCAATGACACCGGACGCTACTTCTGCTGGTCCCAATGACCAAAACAATGTTACCA 1522
 QY 1574 CCATGTGAGGTTCCATCGTCCAGCTGCCACACCTCAGCAACAGCACCGCCGAC 1631
 Db 1523 TCATGGCTAACCTGAAGGTTAAAGATGCAACTCAGATCACTCAGGGGCCCCCGACAGC 1580

RESULT 14
 US-08-866-757-1
 ; Sequence 1, Application US/08866757
 ; Patent No. 5858716
 ; GENERAL INFORMATION:
 ; APPLICANT: ELSHOUBAGY, NABIL A
 ; APPLICANT: LI, XIAOTONG
 ; APPLICANT: BERGSM, DERK J
 ; TITLE OF INVENTION: NOVEL 7TM RECEPTOR (H2CAA71)
 ; NUMBER OF SEQUENCES: 4
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: RAINER & PRESTIA

RESULT 13
 US-08-506-296B-13
 ; Sequence 13, Application US/08506296B
 ; Patent No. 6313265
 ; GENERAL INFORMATION:
 ; APPLICANT: Phillips, Greg
 ; APPLICANT: Cunningham, Bruce A.
 ; APPLICANT: Crossin, Kathryn L.
 ; TITLE OF INVENTION: NEURITE OUTGROWTH-PROMOTING POLYPEPTIDES
 ; CONTAINING FIBRONECTIN TYPE III REPEATS AND METHODS OF USE
 ; NUMBER OF SEQUENCES: 77
 ; CORRESPONDENCE ADDRESS:

```

; STREET: P.O. BOX 980
; CITY: VALLEY FORGE
; STATE: PA
; COUNTRY: USA
; ZIP: 19482
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/866,757
; FILING DATE: 30-MAY-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: PRESTIA, PAUL F
; REGISTRATION NUMBER: 23,031
; REFERENCE/DOCKET NUMBER: GH-70055
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-407-0700
; TELEFAX: 610-407-0701
; TELE: 846169
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4203 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
; US-08-866-757-1

```

```

; TYPE: DNA
; ORGANISM: HOMO SAPIENS
; US-09-153-593-1
;
; Query Match 1.6%; Score 51.6; DB 4; Length 4203;
; Best Local Similarity 54.1%; Pred. No. 0.033; 89; Indels 0; Gaps 0;
; Matches 105; Conservative 0; Mismatches 89; Indels 0; Gaps 0;
;
; QY 684 CAACATGACGGGGCTGGTGGACCTGACCCCTGTCCAGGAACACCCATCAGCCACATCCAGCC 743
; Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
;
; QY 85 CAATCTGCCACCCTTACAGGGCTGACCCCTGCTCAACAAGAATCTCAAGCATCCCTGA 144
; Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
;
; QY 744 CTTTTCCTTTCTGGACCTCGAGAGCTCCGGCTCCCTCGCATCTTGACAGCAATCGGCTGCC 803
; Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
;
; QY 145 CTTTGCATTTACCAACCTTTCAAGCCTGGTAGTTCTGTCATCTTCAACAATAAAATTAG 204
; Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
;
; QY 804 AAGCCTTGGGAGGACACCCCTCCGGGGCTGTCAACCTGCGAGCACCTTATCGTGAACAA 863
; Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
;
; QY 265 TAATAACTTGGGGG 278
; Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
;
; Search completed: June 8, 2003, 19:06:49
; Job time : 237 secs

```

```

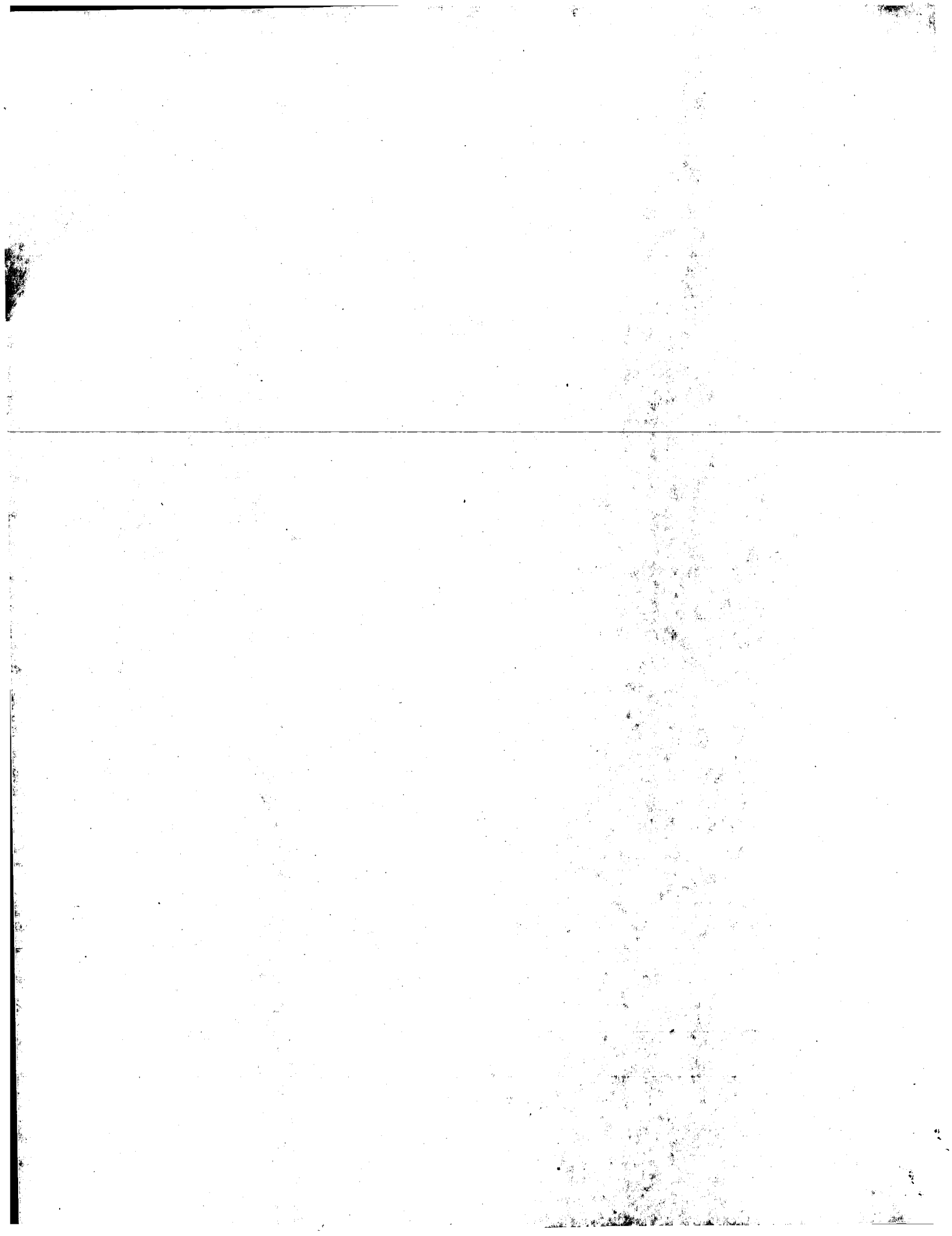
; STREET: P.O. BOX 980
; CITY: VALLEY FORGE
; STATE: PA
; COUNTRY: USA
; ZIP: 19482
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/866,757
; FILING DATE: 30-MAY-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: PRESTIA, PAUL F
; REGISTRATION NUMBER: 23,031
; REFERENCE/DOCKET NUMBER: GH-70055
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-407-0700
; TELEFAX: 610-407-0701
; TELE: 846169
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4203 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
; US-08-866-757-1

```

```

; Query Match 1.6%; Score 51.6; DB 2; Length 4203;
; Best Local Similarity 54.1%; Pred. No. 0.033;
; Matches 105; Conservative 0; Mismatches 89; Indels 0; Gaps 0;
;
; QY 684 CAACATGACGGGGCTGGTGGACCTGACCCCTGTCCAGGAACACCCATCAGCCACATCCAGCC 743
; Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
;
; QY 85 CAATCTGCCACCCTTACAGGGCTGACCCCTGCTCAACAAGAATCTCAAGCATCCCTGA 144
; Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
;
; QY 744 CTTTTCCTTTCTGGACCTCGAGAGCTCCGGCTCCCTCGCATCTTGACAGCAATCGGCTGCC 803
; Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
;
; QY 145 CTTTGCATTTACCAACCTTTCAAGCCTGGTAGTTCTGTCATCTTCAACAATAAAATTAG 204
; Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
;
; QY 804 AAGCCTTGGGAGGACACCCCTCCGGGGCTGTCAACCTGCGAGCACCTTATCGTGAACAA 863
; Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
;
; QY 265 TAATAACTTGGGGG 278
; Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
;
; RESULT 15
; US-09-153-593-1
; Sequence 1, Application US/09153593A
; Patent No. 6174994
; GENERAL INFORMATION:
; APPLICANT: ELSHOUBAGY, NABIL A
; APPLICANT: LI, XIAOTONG
; APPLICANT: BERGSMAN, DEREK J
; TITLE OF INVENTION: NOVEL 7TM RECEPTOR (H2CAA71)
; FILE REFERENCE: GH-70055-1
; CURRENT APPLICATION NUMBER: US/09/153,593A
; EARLIER FILING DATE: 1998-09-15
; EARLIER FILING DATE: 08/866,757
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 4203

```



GenCore version 5.1.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 8, 2003, 18:21:58 ; Search time 411 Seconds
(without alignments)
10705.950 Million cell updates/sec

Title: US-09-831-846-1
Perfect score: 3144
Sequence: 1 gcttggtccctctctcgtga.....ggcggttgattgcttatg 3144

Scoring table: IDENTITY_NUC
Gapop 10_0 , Gapext 1.0
Searched: 870385 seqs, 699768693 residues
Total number of hits satisfying chosen parameters: 1740770

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : Published Applications_NA:
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:
2: /cgn2_6/ptodata/2/pubpna/FCI_NEW_PUB.seq:
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:
6: /cgn2_6/ptodata/2/pubpna/FCIUS_PUBCOMB.seq:
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:
9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:
11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:
12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:
14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Table with columns: Result No., Query Match %, Score, Length, DB ID, Description. Contains 19 rows of search results.

Table with columns: 20-45, 621.8, 19.8, 2855, 9, US-10-243-409-99, Sequence 99, Appl. Contains alignment data for various sequences.

ALIGNMENTS

RESULT 1
US-09-815-626-3
Sequence 3, Application US/09815626
Patent No. US20020076752A1
GENERAL INFORMATION:
APPLICANT: Glucksmann, Maria Alexandra
TITLE OF INVENTION: 33395, A NOVEL HUMAN LEUCINE-RICH REPEAT
FILE OF INVENTION: FAMILY MEMBER AND USES THEREOF
FILE REFERENCE: 10448-031001
CURRENT APPLICATION NUMBER: US/09/815,626
CURRENT FILING DATE: 2001-03-23
PRIOR APPLICATION NUMBER: US 60/191,863
NUMBER OF SEQ ID NOS: 17
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 3
LENGTH: 1887
TYPE: DNA
ORGANISM: Homo sapiens
US-09-815-626-3

Table with columns: Query Match, Best Local Similarity, Mismatches, Indels, Gaps. Contains summary statistics for the alignment.

Db 1375 ATCCAGTACACAGCTCGGCTGATGACATCTCGTCTACAGGATGATCCCGCGGAGAGC 1434
 Qy 1885 AAGGCTTCGTGGTCAACAACCTGGTTCAGGACTGGCTAGACTTGTGTGTGTGCTGCCC 1944
 Db 1435 CGTCTGTTCTCTACAGGACCTGGGCTCAGGCCGAGCCTACGATCTGTGGTGTCTGCCC 1494
 Qy 1945 ATGTGGATGACAGCCACACACTCACGGCCACCAACATCGTGGTGTGGTGTGGTGTGGT 2004
 Db 1495 GTGTATGAGGAGTGGACAGGGGCTCAGGCCACCGGGCTGTGGGCTGTGGGCTGTGG 1554
 Qy 2005 TTCACCAAGGCTGACTACCCGAGTGCAGTTCATGACAGCCAGATTCAGGGCGGACCC 2064
 Db 1555 TCCACCAACCTGGCTGGGCTCAGGGGCGCCGCGCCGCTCCCTTCTGTGGGCGGACG 1614
 Qy 2065 ATGATCTGTGATCGGGGCTATCTGCGGCGCATCTGCGGCGCCGCGCCGCTCCCTT 2124
 Db 1615 ATGATCATCGGCTGGGCGGCTATCGTAGCTCGTACTGGTACTGGTCTTCACTTCGTG 1674
 Qy 2125 ATGGTGGCTACAAAGGTCTGCAACACAGAGGCCCCCAG-----CAAGATGGAGGGCC 2178
 Db 1675 CTAATGGCTTACAAGGTGCAAGGCGGCGCCAGCCCCCGCAAGGCAAGATTCGCCG 1734
 Qy 2179 GTGAGCAATGTACTCCGAGACCAAGGCGGCGCCAGCCCGCTCCAAAGCAGCGCACCA 2238
 Db 1735 GTTAGCAGCGTTTGTCTCCAGACCAAGGCGGCGCTGGGCGCCACGCCCGCCCGC 1794
 Qy 2239 GCCGGGCGCCCGCGCAGG 2257
 Db 1795 CCCCGCGGAGCCCGGG 1813
 RESULT 2
 US-09-815-626-1
 ; Sequence 1, Application US/09815626
 ; Patent No. US20020076752A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Glucksmann, Maria Alexandra
 ; TITLE OF INVENTION: 33395, A NOVEL HUMAN LEUCINE-RICH REPEAT
 ; TITLE OF INVENTION: FAMILY INVENTION;
 ; FILE REFERENCE: 10448-031001
 ; CURRENT APPLICATION NUMBER: US/09/815,626
 ; CURRENT FILING DATE: 2001-03-23
 ; PRIOR APPLICATION NUMBER: US 60/191,863
 ; PRIOR FILING DATE: 2000-03-24
 ; NUMBER OF SEQ ID NOS: 17
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 1
 ; LENGTH: 2558
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (435)...(2318)
 US-09-815-626-1
 Query Match 20.1%; Score 633; DB 10; Length 2558;
 Best Local Similarity 61.3%; Pred. No. 9,9e-166;
 Matches 1078; Conservative 0; Mismatches 665; Indels 16; Gaps 3;

Db 303 CGCCTTCCGCGACCTTGGGGCCCTCGTGGCCCTGACCTGATGGCAACCGCTGACCTC 362
 Qy 807 CCTTGGGAGGACACCCCTCGGGGCTGGTCAACCTGAGCAGCATTATCGTGAACAACA 866
 Db 363 ACTGGGAGGCGCAGCTGGCGGCTGGTCAACTTGGCCACCTCATCTCAGACAACA 422
 Qy 867 CCAGCTGGGCGGATCGCAGATGAGCTTTGAGGACTTCTGCTGACATTTGGAGGATCT 926
 Db 423 CCAGCTGGGAGGCTGGCGGCGGGCCCTGATGATTTGCGCCAGACACTGGAGACT 482
 Qy 927 GGACCTCTCTAACAACAACCTCATGGCCCTGGCTGGACTCCGTCGAGCAGTGGTCAA 986
 Db 483 CGACCTCTCTAACAACAACCTCGAGCAGTGGCCCTGGAGGCGCTGGCGCTGGGCAA 542
 Qy 987 CTTCCACAGCTGAGCCCTGGAACAACCTGCTGGATCACATGCGGAGGACCTTTGC 1046
 Db 543 CGTCAACAACGTTGGGCTCGACACAACCTGCTGGCTTCTGTGCCCGCGGCTTTTC 602
 Qy 1047 AGACTCGAAGAACTGGCCCGCTGGATCTCACTCCAATCGGCTGCAAGACTGCCCC 1106
 Db 603 CGGCTGCAACAGTGGCCCGGCTGGACATGACTCCAACCGCTGACCAACAATCCACC 662
 Qy 1107 TGATCCCACTTTGCCCGCTCCAGGCTTGGCTTTGACAGCCACACCTTTGCCCCAC 1166
 Db 663 CGACCCACTTTTCCCGCTGCTCCCTGCTCGCCAGGCGCCCGGGCTCGCCGCTTCG 722
 Qy 1167 CTTGTCTTTAGTTTGGGGTAAACCACTTCACTGCAATTTGAGCTTCTCTGGGTGG 1226
 Db 723 CCTGGTGTGGCCCTTTGGCGGAACCCCTGCACTGCACTGGAGTGGTGGTGGTGG 782
 Qy 1227 GAGCTCGAGCGGACGATCACTGGAAACCTGFGGCTCCCAAGGGGCTCAAGGGTGG 1286
 Db 783 TCGCTGGCGGGAGGACGACTCGAGGCTCGCGCTCCCACTGCTCTGGCGGCGCG 842
 Qy 1287 CTACTTTCGATGTGCTGAGGAGGAGTTGTGTGGAGCGGCTCTCATACCCAGCA 1346
 Db 843 CTACTTTGGCGGTGGCGGAGGAGTTGTGTGGAGCGGCTCTCATACCCAGCA 902
 Qy 1347 CACACACAAGTTGCTGTGTTCTGGAGGCGGAGGCGGCGCACACTCAAGTGAAGCCATTG 1406
 Db 903 CTCACACCTTGGCTGTGCGCGCAGTGGCGGCTGCTGCGCTGCGCGGCGAGTGG 962
 Qy 1407 GGACCCAGCCCTTATCCACTGGTAGCCCGGAGTGGAGCGCTGAGGAACTCCTC 1466
 Db 963 GGACCCAGGCGGCTGTGCTGTGGTGTCAACCCAGGCGGCTGTGAGCAACTCAAG 1022
 Qy 1467 AAGGACCGTGTATGACAATGGCACCTGGACATCTTATCAACCAATCTCAGGACAG 1526
 Db 1023 CCGTCCCGGCTTCCCAATGGGAGCTGGAGCTGGAGTGTGCTCAGGAGCGGCTGATGG 1082
 Qy 1527 TGGTGCCTTACCTGATGCTGCAATGCTGCGGAGGCGGCGGCGGCGGCGGAGGT 1586
 Db 1083 TGGCATTTTACCTGATTTGGCCCAATGCACTGGGAGGCGGCGGCGGCGGAGTCT 1142
 Qy 1587 CTCATGCTCAGCTGCCACACTCAGCAACAGCACAGCCAGCCGCA--CTGCACCCCGCAAG 1644
 Db 1143 GACTGTGGGTCCCGACCACTTCTCAGCTAGCCAAAGCAGCTGTGACCCCGCGC 1202
 Qy 1645 TCCCGCTCTCAGACATCACTGGCTCCAGCAAGACAGCCGGGAGGTTGGAGGAGTGG 1704
 Db 1203 GGACGGGATCTGTATGCTCTACACCCACCTCCGCTGCTGCTTCTGCAAGGTTGG 1262
 Qy 1705 GCGGAGAGCTCCAAAGCCCCCGGAAACCGGCTGCTGTTGTGCTGAGTACCACC 1764
 Db 1263 CGACACTGGGCCCCCTA-----CGACCTGGGCTCCAGTACTGAGCAGCGGGCC 1314
 Qy 1765 ACCTCGGCGCTGTCAGTGGTGTGTGCAAGTACAGCAGCCCGGTTGAATGATACCAG 1824
 Db 1315 ACAGTGTCTTGTTCAGTGGCGGATCAGCGGCTATCCCGGCAATCCGCACTGACAG 1374
 Qy 1825 CTGAGTACAACCTCTGACGATGAGTACTGATTTACAGGATGATCCAGCTCCCAAC 1884

Db 677. CATGACAGCCCTGCTGCAATCTGAGCCCTGTCGGGAAACACCAATCCGCCACAGTGGCTGCCGG 736
 Qy 747 TTCTTTTGGACCTCGAGAGCTCCGGTCCCTGCAATCTTGACAGCAATCGGCTCCCAAG 806
 Db 737 CGCCTTCGCGACCTCGGGCCCTGGCTGCTCCCTGACCTGGATGGAAACGGCTGACCTC 796
 Qy 807 CTTGGGAGGACACCTTCGGGGCCCTGGTAACTGTCAGACCTTATCGTGAAACAANA 866
 Db 797 ACTGGGAGGGCCAGCTGGGGCCCTGGTAACTTGGCCACCTCATCTCAGCAACA 856
 Qy 867 CGAGCTGGGGCAGCTCGAGATGAGCTTTTGGAGACTTCTGCTGACATGGAGACT 926
 Db 857 CGAGCTGGAGGCTGGGGCCCTGGATGATTTGGCCGAGACACTGGAGACT 916
 Qy 927 GGACCTCTCTACAACAACCTCCATGGCTGCGGTGGACTCCGTCGACGATGGTCAA 986
 Db 917 CGACCTCTCTACAACAACCTCGAGAGCTGCTCCCTGGAGGCCCTGGGGCTGGCAA 976
 Qy 987 CCTCACAGCTGAGCTGGACCAACCTGCTGGATCACATCGCCGAGGACACCTTTGC 1046
 Db 977 CGTAAACAGTTGGCCCTCGACCAACCTGCTGGCTTCTGTCGCCGGGGCTTTTTT 1036
 Qy 1047 AGACCTGAGAACTGGCCGCTGGATCTCACTCCCAATCGGCTGCAGAACTGCCCC 1106
 Db 1037 CGGCTGCAACAGCTGGCCGCTGGATGATGATGATGATGATGATGATGATGATGATGAT 1096
 Qy 1107 TGATCCCATTTTGGCCGCTCCAGGCTTTGGCTTTGACAGCCACACCCCTTTGCCCCACC 1166
 Db 1097 CGACCCACTTCTCCGCTGCTCCCTGCTCGCCAGGCCCGGGCTCGCCGCTCTGC 1156
 Qy 1167 CTTGCTTTTGGGGTAAACCACTTCACTGCAATTTGAGCTTCTTGGCTGG 1226
 Db 1157 CTTGCTGCTGCTTTTGGGGAAACCCCTGCACTGCAACTGCGAGCTGCTGGCTGG 1216
 Qy 1227 GAGCTGAGGGAGTACTGGAACCTGCTGCTCCCGGCTCCCGGGGGCTCAAGGCTG 1286
 Db 1217 TCGCTGGGGAGGAGCACTCGAGCCCTGCGGCTGCGCCCTCCCACTGCTGGGGGGCG 1276
 Qy 1287 CTACTTCTGGCATGCTGAGGAGGTTTGTGCGAGCCGCTCTCATCACCACAGCA 1346
 Db 1277 CTACTTTGGGGTGGGGAGGAGGTTTGTGCGAGCCGCTGCTGCTGGTACTCAGC 1336
 Qy 1347 CACACAAAGTTGCTGGTTCGGAGGCGCAGCGCCACACTCAAGTGCAAAGCCATTGG 1406
 Db 1337 CTACCACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1396
 Qy 1407 GGACCCAGCCCTTATCACTGGTAGCCCGGAGTACCCGCTGCTGCTGCTGCTGCTGCT 1466
 Db 1397 GGACCCAGAGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1456
 Qy 1467 AAGGACCGCTGCTATGACAATGGCACCTGGACATCTTCACTCACACACTCAGACAG 1526
 Db 1457 CGTGGCCGCTTCCCAATGGACGCTGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTG 1516
 Qy 1527 TGGTCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1586
 Db 1517 TGGCATCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1576
 Qy 1587 CTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1644
 Db 1577 GACTGTTGGTCCCAACCACTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1636
 Qy 1645 TCCCGCTCTCAGACATCACTGGCTCCAGCAAGACCCAGCCGGGAGTGGAGGAGTGG 1704
 Db 1637 GGACGGGATCTGATGCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1696
 Qy 1705 GGGGAGAGCTTCCAAAAGCCCCGGAAACGGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1764
 Db 1697 CGACACTGGGGCCCTA-----CCGACCTGGGGTCCAGTGAAGTGAAGTGAAGTGA 1748
 Qy 1765 ACCTGGCCCTGCTCAAGTGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1824
 Db 1749 ACAGTGTCTTGTCCAGTGGGGATCAGCGGCTTATCCCGGGCATCCCGGATGATACCAG 1808

Qy 1825 CTGACGTAACTGCTTGACAGTGAAGTACTGATTTTACAGAGTATCCAGCTCCCAAC 1884
 Db 1809 ATCCAGTAAACAGCTCGCTGATGACATCTCGTCTACAGATGATCCCGCGGAGAGC 1868
 Qy 1885 AAGCCCTTCTGTTCAACAACCTGGTGTGAGGACTGCTAGGACTTGTGTGTGTGCTG 1944
 Db 1869 CGCTGTTCTGCTGACGGACTGGGCTGAGGGGACCTACGATCTGTGGGTGTCTCG 1928
 Qy 1945 ATGTGGATGACACAGCCAGCACACTCACGGCCACAAACATCGTGGGCTGGCCCGAGTTC 2004
 Db 1929 GTGATGAGGAGTGTCCACGGGCTCACGGCCACGGGCTGTGGGCTTGGCCCGCTTC 1988
 Qy 2005 TTCAACAAAGTGTACTCCCGAGTCCAGTCCATGCAACAGCCAGATTTTGGGGCGCAC 2064
 Db 1989 TCCACCGAACCTGCGCTCGGCCATGCGGGCGCCAGCAGCTCCCTTCTTGGCGCGCAG 2048
 Qy 2065 ATGATCTGCTCATCGGGGCTCATCTGTCGCGGCTGCTGCTGCTGCTGCTGCTGCTG 2124
 Db 2049 ATGATCTGCGGCTGGGGGCTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2108
 Qy 2125 ATGTGCTCTACAAGTCTGCAACCCAGGCGCCCGCCAG-----CAAGATGCGAGCGCC 2178
 Db 2109 CTNATGCTCTACAGTGTCCAGCGGCTGCGCGGCTGCGCGGCGCCAGGCAAGATTC 2168
 Qy 2179 GTGACAAATGTTACTCGCAGACCAACCGGCGCCAGCCAGCCGCTTCCAAAGCAGCG 2238
 Db 2169 GTTAGCAGCTTGTCTCCAGCAACCGGCGCCCTGGGCGCCAGCCAGCCCGCCCG 2228
 Qy 2239 GCCGGGCGCCCGCCAGG 2257
 Db 2229 CCGCGCCGAGCCCGCGG 2247

RESULT 3

US-10-245-103-99
 ; Sequence 99, Application US/10245103
 ; Publication No. US20030068778A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Baker, Kevin
 ; APPLICANT: Eaton, Dan
 ; APPLICANT: Filvaroff, Ellen
 ; APPLICANT: Goddard, Audrey
 ; APPLICANT: Grimaldi, J. Christopher
 ; APPLICANT: Gurney, Austin
 ; APPLICANT: Smith, Victoria
 ; APPLICANT: Stephan, Jean-Phillippe
 ; APPLICANT: Watanabe, Collin
 ; APPLICANT: Wood, William
 ; APPLICANT: Zhang, Zemin
 ; APPLICANT: Fong, Sherman
 ; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
 ; FILE REFERENCE: P3630R1C112
 ; CURRENT APPLICATION NUMBER: US/10/245,103
 ; CURRENT FILING DATE: 2002-09-17
 ; PRIORITY APPLICATION NUMBER: 10/197942
 ; PRIORITY FILING DATE: 2002-07-18
 ; PRIORITY APPLICATION NUMBER: 60/059114
 ; PRIORITY FILING DATE: 1997-09-17
 ; PRIORITY APPLICATION NUMBER: 60/063046
 ; PRIORITY FILING DATE: 1997-10-24
 ; PRIORITY APPLICATION NUMBER: 60/065027
 ; PRIORITY FILING DATE: 1997-11-10
 ; PRIORITY APPLICATION NUMBER: 60/079689
 ; PRIORITY FILING DATE: 1998-03-27
 ; PRIORITY APPLICATION NUMBER: 60/086478
 ; PRIORITY FILING DATE: 1998-05-22
 ; PRIORITY APPLICATION NUMBER: 60/087607
 ; PRIORITY FILING DATE: 1998-06-02
 ; PRIORITY APPLICATION NUMBER: 60/089801
 ; PRIORITY FILING DATE: 1998-06-18
 ; PRIORITY APPLICATION NUMBER: 60/090557

1635 CTACACACCTCTGGCTGTGCCCGCAGGTCCGCCGGCTCCCGGGCAGTGGG 1694
 1407 GGACCCAGCCCCCTTATCCACTGGTAGCCCGCCAGCCCTGGTAGGAACTCCTC 1466
 1695 GGACCCAGCCCCCGTGGTGGTGTCCAGCCCGGGCTGCTAGGCACTCAAG 1754
 1467 AAGGACCGGTGTATGACAAATGGCACCTTTCATCACCAATCTCAGACAG 1526
 1755 CGGTGCCCGGGCTTCCCAATGGGAGCGTGGAGTGTGGTCCAGGACCGGGT 1814
 1527 TGGTGCCTTCACTGCAATGCTCCGAGAGGAGCCAGGGCCATGGTGGAGGT 1586
 1815 TGGCACTTCACTGCAATGCTCCGAGAGGAGCCAGGGCCATGGTGGAGGT 1874
 1587 CTCCATGCTCAGCTCCACACCTCTCAGCTAGCCAAACAGCACAGCTGTG 1644
 1875 GACTGTGGGTCCCAACACCTCTCAGCTAGCCAAACAGCACAGCTGTG 1934
 1645 TCCCGCCTCTCAGACATCACTGGCTCCAGCAAGACCAGCCGGGGAGGTGG 1704
 1935 GGAGGGGATCTCTGATGCTCTCACCCACCTCCGCTGCTTGTGCTGCAAGT 1994
 1705 GCGGAGAGCTCCAAAGCCCGCCGGAACGGGCTGTGCTTGTCTGAAAGTGA 1764
 1995 CGACACTGGGCCCCCTA-----CGACCGTGGGCTCCAGTGTGACGACGG 2046
 1765 ACCTGGCCCTGGTCAAGTGGTGTGTAGCAAGTCAAGACCCCGGGTGAAG 1824
 2047 ACAGTCTCTTGTTCAGTGGCCGGATCAGCGGCTATCCCGGCACTCCG 2106
 1825 CTGAGTAACTCTCTGAGTGTGATGATTTACAGGATGATCCAGCCTCAAC 1884
 2107 ATCCAGTAAACACTCGGCTGATGATCTCTGCTTACAGGATGATCCCGG 2166
 1885 AAGCCCTTGGTCAAAACCTGGTGTAGGGACTGGCTACGACTTGTGTGCTG 1944
 2167 CGCTGTTCTGTGACGGACTGGCTCAGCGGACTACGATCTGTGCTGCTCG 2226
 1945 ATGTGGATGACAGCCAGCACACCTACCGGCCAACAACTCCTGGCTGGCC 2004
 2227 GTGATGAGGACAGCCCGGCTCAGGCCACGGGCTGAGGGCTGGCCCGCT 2286
 2005 TTCACCAAGTACTACCCAGTGCAGTCCATGACAGCCAGATTTCTGGCGG 2064
 2287 TCCACCAAGTACTACCCAGTGCAGTCCATGAGGGGCGCCAGCCTCCCT 2346
 2065 ATGATCTTGGTCACTCGGGGGCATCATGTTGGCCACAGTGTGCTTTCAT 2124
 2347 ATGATCACTGGCTGGGGGGTCACTGTTAGCCTCGTACTGTTCTTTCAT 2406
 2125 ATGTTGGCTTACAGGTGTCAACCCAGGAGCCCGGAG-----CAAGATG 2178
 2407 CTAATGCGCTACAAGGTGCAAGGCGGCGGAGCCCGGCAAGGCAAGATT 2466
 2179 GTGAGCAATGTACTCGCAGACCAACCGGGCCCGCCAGCCCGCTTCCAA 2238
 2467 GTTAGCAGGTTTGTCTCCAGACCAACCGGGCCCTGGGCGCCCGCCAG 2526
 2239 GCCGGGGCCCCCGGAGG 2257
 2527 CCGCCCGGAGCCCGGG 2545

507 GTTTCGGTGTGCGCCCTCCAAAGGGGCTCTTTTTPACCCCTGATATTGAC 566
 798 GTCAGCCACACCCAGCCCAATGTCCTCCCGCGCTGCCGCTCCAGACACAG 857
 567 GGGGACCTGTGCCCTTCCAAAGGGGCTCTTTTTPACCCCTGATATTGAC 626
 858 AAGGTGTCTGTCCAGGGGAGGCTCTCTGTTGCTGCTCCACCTCGCTGG 917
 627 AGTGGAGCTGGCCGAGCACTTCACTATCCATCCACATCAGCCGAGACTT 686
 918 AGCCGAGCTCGGCTGSCAGACAACTTCAATGCGCTCCGTCGCGCCGCG 977
 687 CATGACGGGCTGTGGACCTGACCTCTCCAGGAAACACCATCAGCCCAT 746
 978 CATGACGGGCTGTGATCTGAGCTGTGCGGAAACACCATCAGCCCAT 1037
 747 TTCCTTCTGGAACCTGAGAGCTTCGCTCCCTGATCTTACAGCAATGG 806
 1038 CGCCTTCCCGACCTGCGGGCCCTCGTGCCTGCACTGGATGGCAACCG 1097
 807 CTTGGGAGGACACCTCCGGGCTGTGCAACCTGACCTTATCGTGAACAAC 866
 1098 ACTGGGAGGGCCAGCTGCGGCTGTGCACTGCGGCACTCATCTCAGAAC 1157
 867 CCAGTGGGCGGCTGCGAGATGAGGCTTTGAGGACTTCTGATGAGAGT 926
 1158 CCAGTGGGCGGCTGCGGCGGCGGCTGATGATGTTGCGGAGACCTGG 1217
 927 GGACTTCTTCAACAACTTCCATGGCTGCGGAGTCCGTTGGAGGATGTC 986
 1218 CGACTCTCTTCAACAACTTCCATGGAGCTGCTGCGGCGGCTGCGGCA 1277
 987 CCTCACCAGCTGAGCCTGGACCAACCTCTGATCAATCCCGAGGGCAC 1046
 1278 CGTCAACAGTGGGCTTGCACCAACCTCTGCTGCTGCTGCTGCTGCT 1334
 1047 AGACTGAGAACTGGCCGCTGGATCTCACCTCCATCGGCTGAGAGCTGC 1106
 1335 CGCCCTGACAACTGGCCGCTGGATGATGATGATGATGATGATGATG 1394
 1107 TGATCCACTTTTGGCTCCCGCTCCAGGCTTGGCTTGGACAGCCACCC 1466
 1395 CGACCACTTCTTCCCGCTCCCGCTGCTGCGAGGCGCCCGGCTGCTG 1454
 1167 CTTGTCTTGTGTTGGGGTAAACCCTTCACTGCAATTTGAGCTTCTTGG 1226
 1455 CTTGGTGTGGCTTGGCGGAAACCCCTGCACTGCACTGCGAGCTGGT 1514
 1227 GAGGCTGAGCGGAGGATGACCTTGAAACCTGTGGCTCCCGAGGGGCTCA 1286
 1515 TCGCCCTGGCGGGAGGAGACTGAGGCTTGGCGTCCCACTGCTTGGGG 1574
 1287 CTACTTCTGGCATGTCCGCTGAGGAGGTTTGTGCGAGCCCTCTCATC 1346
 1575 CTACTTCTGGCGGTGGCGGAGGAGGTTTGTGCGAGCCCGGCTGGTGA 1634
 1347 CACACAAAGTCTGCTGTTCTGGAGGCGGAGGCGGACACTCAAGTGA 1406

RESULT 4
 US-10-245-107-99
 ; Sequence 99, Application US/10245107
 ; Publication No. US2003068779A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Baker, Kevin
 ; APPLICANT: Eaton, Dan
 ; APPLICANT: Filvaroff, Ellen
 ; APPLICANT: Goddard, Audrey
 ; APPLICANT: Grimaldi, J. Christopher

PRIOR FILING DATE: 1998-06-24
 PRIOR APPLICATION NUMBER: 60/090689
 PRIOR FILING DATE: 1998-06-25
 Remaining Prior Application data removed - See File Wrapper or PALM.
 NUMBER OF SEQ ID NOS: 116
 SEQ ID NO 99
 LENGTH: 2855
 TYPE: DNA
 ORGANISM: Homo Sapien
 US-10-245-103-99

Query Match 19.8%; Score 621.8; DB 9; Length 2855;
 Best Local Similarity 61.3%; Pred. No. 1.3e-162;
 Matches 1078; Conservative 0; Mismatches 662; Indels 19; Gaps 4;

APPLICANT: Gurney,Austin
 APPLICANT: Smith,Victoria
 APPLICANT: Stephan,Jean-Phillippe
 APPLICANT: Watanbe,Colin
 APPLICANT: Wood,William
 APPLICANT: Zhang,Zemin
 APPLICANT: Fong,Sherman
 TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
 TITLE OF INVENTION: ACIDS ENCODING THE SAME
 FILE REFERENCE: P3630R1C71
 CURRENT APPLICATION NUMBER: US/10/245,107
 CURRENT FILING DATE: 2002-09-16
 PRIOR APPLICATION NUMBER: 10/197942
 PRIOR FILING DATE: 2002-07-18
 PRIOR APPLICATION NUMBER: 60/059114
 PRIOR FILING DATE: 1997-09-17
 PRIOR APPLICATION NUMBER: 60/063046
 PRIOR FILING DATE: 1997-10-24
 PRIOR APPLICATION NUMBER: 60/065027
 PRIOR FILING DATE: 1997-11-10
 PRIOR APPLICATION NUMBER: 60/079689
 PRIOR FILING DATE: 1998-03-27
 PRIOR APPLICATION NUMBER: 60/086478
 PRIOR FILING DATE: 1998-05-22
 PRIOR APPLICATION NUMBER: 60/087607
 PRIOR FILING DATE: 1998-06-02
 PRIOR APPLICATION NUMBER: 60/089801
 PRIOR FILING DATE: 1998-06-18
 PRIOR APPLICATION NUMBER: 60/090557
 PRIOR FILING DATE: 1998-06-24
 PRIOR APPLICATION NUMBER: 60/090689
 PRIOR FILING DATE: 1998-06-25
 Remaining prior Application data removed - See File Wrapper or PALM.
 NUMBER OF SEQ ID NOS: 116
 SEQ ID NO 99
 LENGTH: 2855
 TYPE: DNA
 ORGANISM: Homo Sapien
 US-10-245-107-99

Query Match 19.8%; Score 621.8; DB 9; Length 2855;
 Best Local Similarity 61.3%; Pred. No. 1.3e-162;
 Matches 1078; Conservative 0; Mismatches 662; Indels 19; Gaps 4;

QY	507	GTGTCGGTGTGCGACCGCTGCCCAAGTACTGTGTGCCAGAACTCTCTGAGTCACT	566
Db	798	GTACGCCACACCCAGCCCATGTCCCGCGCTGCCGCTGCCAGACACAGTCTGCTGCCCT	857
QY	567	GGGGACCTGTGCCCTCCAAAGGGGCTGCTTTTTPACCCCTGATATTGACCCGGCGGAC	626
Db	858	AAGCGTGTGTGCCAGGGGAGGCTCTCTGTTGCGCACCTCGCTGGACCCGGCGGGC	917
QY	627	AGTGGAGCTGGCTGGCGGCACTTATATCCATCCATCAGCCCGCAGGACTTTGCCAA	686
Db	918	AGCCGAGCTGCGGCTGGCAGACAACTTATCGCTCCGTGCGCGCGCGGACCTGGCCAA	977
QY	687	CATGAGGGGCTGTGGACTGACCTGTCCAGGAACACATCAGCCACATCCAGCCCTT	746
Db	978	CATGACAGCCCTGTGCATCTGAGCTGTGCGGAAACACATCCCGCCACGTGGCTGCCGG	1037
QY	747	TTCTTTTCTGGACCTCGAGACCTCCGCTCCCTGCACTTTGACAGCAATCGGCTGCCAAG	806
Db	1038	CGCCTTTCGGGACCTCGGGCCCTCGTGGCTGACCTGATGGCAACCGGCTGACCTC	1097
QY	807	CCTTGGGAGGACACCTCGGGCCCTGGTCAACCTGCGAGCCTTATCGTGAACAA	866
Db	1098	ACTGGCGAGGGCCAGCTGCGGGCCCTGCTCACTTGGCCACCTCATCTCCAGCAAA	1157
QY	867	CCAGCTGGGCGGACCTCGAGATGAGGCTTTTGGAGACTTCCCTGCACATTTGAGGACT	926
Db	1158	CNAGCTGGAGGCGCTGGCGGCGCCCTGGATGATTTGTCGGGAGACACTGGAGACCT	1217
QY	927	GGACCTCTCTACAACAACCTCCATGCGCTGCCGTGGGACTCCGTCGGACGATGGTCAA	986

1218 CGACCTCTCTACAACAACCTCGAGACCTCCCTGGGAGGCCCTTGGCCCGCTGGCAA 1277
 987 CCTCCACAGCTGAGCCTGGACCAAACTGTGTGATCACATGCGCGAGGGACCTTTGC 1046
 1278 CGTNAACAGTTGGCCCTCGAACCAACTGCTGCTTCTGTGCC--GGGCTTTTC 1334
 1047 AGACCTCGAGAACTGGCCCGCTGGATCTCACCTTCAATCGGTGCAAGAGCTCCGCC 1106
 1335 CGCGCTGCAAGCTGGCCCGCTGGACATGACCTTCAACCGCTGACCAAACTCCACC 1394
 1107 TGATCCATCTTTGCCCGCTCCAGCTTCCGCTTTGACAGCCACACCTTTGGCCACC 1166
 1395 CGACCCACTTCTCCCGCTCCCTCGCCAGGACCCCGGGCTCCCGCCCTCTGC 1454
 1167 CTTGCTTCTTGGGGTAAACCACTTCACTGCAATTTGAGCTTCTCTGGCTGG 1226
 1455 CTTGCTTGGCTTGGGGAAACCCCTGCACTGCAACTGCGAGCTGCTGCGGCTGG 1514
 1227 GAGGCTCGAGCGGACGATGACCTGGAAAACCTGTGGCTCCCGAGGGGCTTCAAGGGTGG 1286
 1515 TCGCTGGCGGGAGGACGACCTCGAGGCTGCGCTCCCACTGCTTGGCGGCGC 1574
 1287 CTACTTCTGGCATGTGCGTGGAGGAGTGTGTGCGAGCGCCCTCTCATCACCAGCA 1346
 1575 CTACTTCTGGCGGTGGCGGAGGAGTGTGTGCGAGCGCCCTGGTGTGACTCACCG 1634
 1347 CACACACAAGTTGCTGTTCTGGAGGGCAGCGGCCACACTCAAGTGCAAAAGCATGG 1406
 1635 CTCACCACTTGGCTGTCGCCGAGTGGCGGCTGCCCTGCGCTGCGGGGAGTGG 1694
 1407 GGACCCAGCCCTTATCCACTGGGTAGCCCCCGATGACCCCTGGTAGGAACTCCTC 1466
 1695 GGACCCAGAGCCCGTGTGGTGTGACCCCGAGGGCGGCTGAGGCAACTCAAG 1754
 1467 AAGGACCGCTGTATGACAAATGGCACTTGGACATTTTATCACCACATCTCAGGACAG 1526
 1755 CCGTCCCGCGCTTCCCAATGGGACGCTGGAGCTGTGTTTCCAGGACCGGGTATGG 1814
 1527 TGGTCCCTTCACTGCTGCAATGCTGCGGAGAGGCGCCAGGCTGGTGGAGT 1586
 1815 TGGCATTTCACTGCTGCGGCAATGAGCTGGGAGGCGCCAGCTGCTGGAGCT 1874
 1587 TCCATCTGTCAGCTGCCACACTCAGCAACAGCAGCCGCA--CTGACCCCGCAAG 1644
 1875 GACTGTGGTCCCGCCACACTCTCTCAGCTAGCCAAACAGCAGCTGTGACCCCGCG 1934
 1645 TCCCGCTCTAGACATCACTGGTCCAGCAAGACAGCCGGGAGGTTGGAGCGTGG 1704
 1935 GGACGGGATCTGATGCTCTACCCCACTCCGCTGCTGCTTCTGCAAGGTGGC 1994
 1705 GCGGAGAGCTCCCAAAAGCCCCCGGAAAGGGCTGTGTTGCTGCTGCAAGTACCAC 1764
 1995 CGACACTGGGCCCCCTA-----CCGACCGTGGCTGCGAGTACTGAGCAGCCGGGGC 2046
 1765 ACCTCGGCTGCTCAAGTGTCTGTGACAGTGTGACACCCCGGCTGAAGATGACCAAG 1824
 2047 ACAGCTGCTTGTCCAGTGGCGGATCAGCGGCTATCCCGGCTATCCCGGCTATACCAG 2106
 1825 CTGAGTACAATGCTCTGACGATGAGGTACTGATTTTACAGGATGATCCAGCCTTCAAAC 1884
 2107 ATCCAGTACAACAGCTCGCTGATGACATCTCTGCTACAGGATGATCCCGGCGGAGAC 2166
 1885 AAGGCTTGGTGTCAACAACTGGTGTGAGGACTGGCTTACGACTTGTGTGCTGGCC 1944
 2167 CGCTGTTCTGCTCAGGACCTGGGCTGAGCGGCACTACGATCTGTGCGTGTGCTGCC 2226
 1945 ATGTTGGATGACACAGCCACACACTCAGCGCCACCAACATCTGGGCTGCCCGCAGTTTC 2004
 2227 GTGTATGAGACAGCCCGGCTCAGGGCCACCGGCTGTGGGCTGCGCCCGCTTC 2286
 2005 TTCACCAAGGCTGACTACCCGAGTGGCAGTCCATTCACAGCCAGTTCGTGGGCGGACCC 2064

507 GTTTCGGTTCGACGGCTCCCAAGTACTGTGTCTGCCAGAACTGTCTGATGACT 566
 798 GTACGCCACCCAGCCCATGTCGCCGGCTGCCGCTGCCAGACAGATCCCTGCCCT 857
 567 GGGGACCTGTGCCCTCCAAAGGGCTGCTTTGTACCCCTGATATGACCCGGGAC 626
 858 AAGCGTGTGTGCCAGGGGACGGCTCTGTTCGTGCCACCTCGCTGGACCCGGCC 917
 627 AGTGGAGCTGCCCTGGGGCAACTTCATCATCCACATCAGCCGCCAGGACTTTGGCAA 686
 918 AGCCGAGCTGCCGCTGGCAGCAACTTCATGCCCTCCGTCGGCCGCCGACCTGCCAA 977
 687 CATGACGGGCTGGTGGACCTGACCTGTCAGGAACACCATCAGCCACATCCAGCCCTT 746
 978 CATGACAGGCTGCTGCACTGAGCCTGTGCGGGAACACCATCCGCACTGGCTGCCGG 1037
 747 TTCCTTTCGACCTCGAGGCTCCGCTCCCTGGATCTTGACAGCAATCGGCTGCCAAG 806
 1038 CGCCTTCGCCGACCTGGGGCCCTGGCTGCCCTGACCTGGATGGCAACCGGCTGACCTC 1097
 807 CCTTGGGAGGACACCCCTCCGGGGCCTGGTCAACCTGCAGACACTTATCGTGAACAACA 866
 1098 ACTGGGAGGGCAGCTGCCGGCCTGGTCAACTTGGCCACCTCATCTCAGCAACA 1157
 867 CCAGCTGGGGCATCGAGATGAGGCTTTGAGACTTCTGCTGACATTTGAGGATCT 926
 1158 CCAGCTGGCAGGCTGGCGCCCTGGATGATTTGTCGAGACACTGGAGGACCT 1217
 927 GGACTCTCTACAAACCTCCATGCTCCGCTGGGACTCCGTCGGAGCATGGTCAA 986
 1218 CGACTCTCTTAAACAACCTCGAGGACTGCCCTGGGAGGCTCCGGCCCTGGGCAA 1277
 987 CCTCCACAGCTGAGCTGGACCAACAACCTGCTGGATCACATGCCGAGGSCACTTTGC 1046
 1278 CGTCAACACGTTGGCCCTCGACCAACAACCTGCTGGCTTCTGTGCCCC--GGCGTTTTC 1334
 1047 AGACTCGAATAATGGCCGCTGGATCTCACCTCAATCGGTCGAGAACTGCCCTCC 1106
 1335 CGCCCTGCACAAGCTGGCCCGCTGGACATGACCTCCAAACCGCTGACCAACTCCACC 1394
 1107 TGATCCCACTTTGCCCGCTCCCAGGCTTCCGCTTTGACAGCACAACCTTTGCCCAACC 1166
 1395 CGACCCACTTCTCCCGCTGCCCTGCTCGCAGGCCCCGGGGCTGCCCGCTCTGC 1454
 1167 CTTGTCTTTAGTTTTGGGGTAAACCACTTCACTGCAATTTGAGTCTCTGGCTGC 1226
 1455 CTTGCTGCTGGCTTTGGGGGAACCCCTGCACTGCAACTGCAACTGGAGCTGGTGGCTGCG 1514
 1227 GAGGCTCGAGCGGACGATGACCTGGAAACCTTGGCTCCCGAGGGGCTCAAGGGTCC 1286
 1515 TCGCCTGGCGGGAGGACCTCGAGGCTCCGCTGCCCTCCCACTGCTTGGGGGGCCG 1574
 1287 CTACTTCGGCATGTCGTGGAGGATTTGTGGAGCGCCCTCTCATCACCAGCA 1346
 1575 CTACTTCGGCGGGTGGGAGGATTTGTGGAGCGCCCTGGCTGACTCACC 1634
 1347 CACACAAAGTCTGTTCTGGAGGGCCAGCGGCACTCAAGTCAAAAGCACTGG 1406
 1635 CTACCACTCTGGCTGTGCCCGAGGTGCCCGGTGCTGCCCTGCCGGGAGTGG 1694
 1407 GGACCCAGCCCTTATCCACTGGTAGCCCGATGACCCCTGGTAGGAACTCTCTC 1466
 1695 GGACCCAGCCCGTGTGGTGGTGTACCCAGGGCCGGTGTGACTAGGCACTCAAG 1754
 1467 AAGGACCGCTGTATGACAAATGGCCCTGGACATCTTCAATCACACATCTCAGACAG 1526
 1755 CCGTCCCGGCTTCCCAATGGAGGCTGGAGCTGGTGGTCCAGCCGCGGTGATGG 1814
 1527 TGFTGCTTCACTGCAATTCGCAATGCTGCGGAGGAGCCACCGCCATGTTGGAGGT 1586
 1815 TGGCATCTTCACTGCAATTCGCAATGCTGCGGAGGAGCCACAGCTGCTGTGGAGCT 1874
 1587 CTCCTGCTCAGCTGCCACACCTCAGCAACAGCACCCAGCCGCA--CTGACCCCCCAAG 1644

2287 TCCACCGAACCTCGCGCTGGGCCATCGGGGCGCCAGCGCTCCCTTCTGGGGCGCAGC 2346
 2065 ATGATCCTGGTTCATCGGGGGCATCATCGTGGCCAGCGCTGGTGTTCATCGTTCATCTC 2124
 2347 ATGATCATCGCCTGGCGGCTCATCGTAGCTAGCTGCTGCTGCTTTCATCTTCTGGTGTG 2406
 2125 ATGTGGCTTACAGGCTGCAACCAACGAGCCCGCCAG-----CAAGATGGCAGCGGCC 2178
 2407 CTAATCGCTACAAGGTGACCGGGCCAGCCCGGCGCCAGCCCGGCAAGGCAAGATTCGCGGCT 2466
 2179 GTGAGCAATGTACTCGAGACCAACCGCGCCCGCCAGCCCGCTTCCAAAGCAGCGCACCA 2238
 2467 GTTAGCAGGCTTTGCTCCAGACCAACCGCGCCCTGGGCCCGCCACCGCCCGCCCG 2526
 2239 GCCGGGGCCCCCGCAGG 2257
 2527 CCGGCCCGGAGCCCGG 2545

RESULT 5
 US-10-245-143-99 ; Sequence 99, Application US/10245143
 ; Publication No. US20030088780A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Baker, Kevin
 ; APPLICANT: Eaton, Dan
 ; APPLICANT: Filvaroff, Ellen
 ; APPLICANT: Goddard, Audrey
 ; APPLICANT: Grimaldi, J. Christopher
 ; APPLICANT: Gurney, Austin
 ; APPLICANT: Smith, Victoria
 ; APPLICANT: Stephan, Jean-Phillippe
 ; APPLICANT: Watanabe, Colin
 ; APPLICANT: Wood, William
 ; APPLICANT: Zhang, Zemin
 ; APPLICANT: Fong, Sherman
 ; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
 ; FILE REFERENCE: P3630R1C90
 ; CURRENT APPLICATION NUMBER: 10/197942
 ; CURRENT FILING DATE: 2002-09-16
 ; PRIOR APPLICATION NUMBER: 10/197942
 ; PRIOR FILING DATE: 2002-07-18
 ; PRIOR APPLICATION NUMBER: 60/059114
 ; PRIOR FILING DATE: 1997-09-17
 ; PRIOR APPLICATION NUMBER: 60/063046
 ; PRIOR FILING DATE: 1997-10-24
 ; PRIOR APPLICATION NUMBER: 60/065027
 ; PRIOR FILING DATE: 1997-11-10
 ; PRIOR APPLICATION NUMBER: 60/079689
 ; PRIOR FILING DATE: 1998-03-27
 ; PRIOR APPLICATION NUMBER: 60/086478
 ; PRIOR FILING DATE: 1998-05-22
 ; PRIOR APPLICATION NUMBER: 60/087607
 ; PRIOR FILING DATE: 1998-06-02
 ; PRIOR APPLICATION NUMBER: 60/089801
 ; PRIOR FILING DATE: 1998-06-18
 ; PRIOR APPLICATION NUMBER: 60/090557
 ; PRIOR FILING DATE: 1998-06-24
 ; PRIOR APPLICATION NUMBER: 60/090689
 ; PRIOR FILING DATE: 1998-06-25
 ; Remaining Prior Application data removed - See File Wrapper or PALM.
 ; NUMBER OF SEQ ID NOS: 116
 ; SEQ ID NO 99
 ; LENGTH: 2855
 ; TYPE: DNA
 ; ORGANISM: Homo Sapien
 US-10-245-143-99
 Query Match 19.8%; Score 621.8; DB 9; Length 2855;
 Best Local Similarity 61.3%; Pred. No. 1.3e-162;
 Matches 1078; Conservative 0; Mismatches 662; Indels 19; Gaps 4;

```

; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/063046
; FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/065027
; FILING DATE: 1997-11-10
; PRIOR APPLICATION NUMBER: 60/079689
; FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/086478
; FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/087607
; FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/089801
; FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/090557
; FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090689
; FILING DATE: 1998-06-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 116
; SEQ ID NO 99
; LENGTH: 2855
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-245-771-99
Query Match 19.8%; Score 621.8; DB 9; Length 2855;
Best Local Similarity 61.3%; Pred. No. 1.3e-162;
Matches 1078; Conservative 0; Mismatches 662; Indels 19; Gaps 4;
QY 507 GTTTCCTGCTGACAGCCCTGCCCAAGTACTGCTGTCGCCAAGAAATGCTGTGAGTCACT 566
Db 798 CTCACACACACCCAGCCCATGTCCTCCCGCCGCTGCCAGACACAGTCCGCTGCCCT 857
QY 567 GGGGACCCTGTGCCCTCAAGGGGCTGCTCTTTGTTACCCCTGATATGACCCGGGAC 626
Db 858 AAGCTGTGTGCCAGGGGAGGCTCTGTGTCGCCACCCTCGCTGGACCCGGCGG 917
QY 627 AGTGAGCTGGCCCTGGGGGCAACTTTCATCATCCACATCAGCCGCCAGGATTTGGCCAA 686
Db 918 AGCCAGCTGGGCTGGCAGACAACTTCATCGCTCCGTGGCCGCGGACCTGGCCAA 977
QY 687 CATGACGGGGTGTGGACCTGACCCCTGTCCAGAAACACATCAGCCACATCCAGCCCTT 746
Db 978 CATGACGGCCCTGTGCATCTGAGCCTGTCCGGGAAACACCATCCGCCACCTGGTCCGG 1037
QY 747 TTCCCTTTCTGGACCTCGAGAGCCTCCGCTCCCTGATCTTGACACAATCCGCTGCCAAG 806
Db 1038 CGCCTTCGCCACCTGGGGCCCTCGCTGCTGCACCTGGATGGCAACCCGCTGACCTC 1097
QY 807 CCTTGGGAGGACACCCCTCCGGGGCTGGTCAACCTCGAGCACCTTATCGTGAACAACAA 866
Db 1098 ACTGGCCGAGGGCCAGCTCGGGGCTGGTCAACTTGGCCACCTCATCTCAGCAACAA 1157
QY 867 CCAGCTGGGGGCGATCGCAGATGAGGCTTTTGGAGACTTCTCTGTCGACATGGAGGATCT 926
Db 1158 CCAGCTGGCAGCGCTGGGGCCGCGCCCTGGATGATTTGTCGGAGACACTGGAGACCT 1217
QY 927 GGACCTCTTACAAACACCTTCCATGGGCTGCCGTTGGGACTCCGTTGCGACGATGGTCAA 986
Db 1218 CGACCTCTTACAAACACCTCGAGCAGCTGCCCTGGGAGGGCCCTGGGGCCGCTGGGCAA 1277
QY 987 CCTCCACAGCTGAGCCTGGACCAACACTGCTGGATCACATCGCCGAGGGACCTTTTGC 1046
Db 1278 CGTCAACAGTTGGCCCTGACCAACACTGCTGGCTTCTGTGCC--GGGCTTTTTC 1334
QY 1047 AGACTCGAATACTGGCCCGCTGGATCTCACTCCAAATCGGCTGCGAAGCTGCCCC 1106
Db 1335 CGCCCTGCAAGCTGGCCCGCTGGACATGACCTCCAAACCCCTGACCCACAATCCACC 1394
QY 1107 TGATCCCATTTGGCCCGCTCCAGGCTTCCAGCTTTGACAGCCACACCTTTTGGCCACC 1166
Db 1395 CGACCCACTTCTCCCGCCCTGCCCTCGCCAGGGCCCGGGGCTCGCCGCGCTGTC 1454

```

RESULT 6

```

US-10-245-771-99
; Sequence 99, Application US/10245771
; Publication No. US20030068781A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin
; APPLICANT: Eaton, Dan
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Goddard, Audrey
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Phillippe
; APPLICANT: Watanabe, Colin
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; APPLICANT: Fong, Sherman
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P36301C98
; CURRENT APPLICATION NUMBER: US/10/245, 771
; PRIORITY FILING DATE: 2002-09-16
; PRIOR APPLICATION NUMBER: 10/197942
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/059114

```

QY 1167 CTTGTCCTTTAGTTTGGGGTAAACCACTTCACTGCAATTTGAGCTTCTCTGGCTGG 1226
 Db 1455 CCTGCTGCTGGCCCTTTGGCGGAACCCCTGACATGCAACTGCGAGCTGGTGGCTGG 1514
 QY 1227 GAGGCTCGAGCGGACAGTACCTGGAACCTGTGGCTCCCGAGGGGCTCAAGGGTGG 1286
 Db 1515 TGCCTTGGCGGGAGGAGCACTCGAGGCTTGGCCCTCCCGAGGCTGCTGGGGCGG 1574
 QY 1287 CTACTTCTGGCATGCTGCTGCTGAGGAGAGTTTGTGCGAGCGGCTCATCACCCAGCA 1346
 Db 1575 CTACTTCTGGCGGGTGGCGAGGAGTTTGTGCGAGCGGCTCATCACCCAGCA 1634
 QY 1347 CACACACAAGTTGCTGCTGCTGAGGCGGAGGCGGCACTCAAGTGAAGCAAGCAATGG 1406
 Db 1635 CTACCCACCTCTGGCTGCTGCGGAGGCTGCGCGGCTGCTGCGCGGAGTGGG 1694
 QY 1407 GAGCCCGAGCCCTTTTCACTGGGTAGCCCGGATGACCCGCTGCTGAGGAACTCTC 1466
 Db 1695 GAGCCAGAGCCCGCTGCTGCGGTTGGGTGTACCCAGGCGGCTGTAGGCAACTCAAG 1754
 QY 1467 AAGACCGCTCTATGACAAATGCACTGCACTGCACTTTTCAACACACATCTCAGGACAG 1526
 Db 1755 CCGTGGCCCGGCTTCCCAATGGAGCGCTGGAGCTGTGTACCCAGGCGGCTGTAGG 1814
 QY 1527 TGGTGCCTTCACTGCACTGCTGCAATGCTGCGGAGAGGCAAGGCGGCTGAGGAGT 1586
 Db 1815 TGGCATCTTCACTGCACTGCGGCAATGCGGCAATGCGGAGGCGGCTGCTGAGGAGT 1874
 QY 1587 CTCATCGTCCAGTCCACACCTCAGCAGCAGCAGCCGCA--CTGCACCCCGGCAAG 1644
 Db 1875 GACTGTGGTCCCCACCACTCTCAGTAGCCAAAGCAGCAGCAGTGTGAGGAGTGGG 1704
 QY 1645 TCCCGCTCTCAGACATCACTGGCTCCAGAAAGCAGCAGCAGGCGGAGGAGGAGTGGG 1994
 Db 1935 GGACGGGATCTGATGCTCTACCCCACTCCGCTGCTGCTGCTGCTGCTGCTGCTG 1994
 QY 1705 GCGGAGAGCTCCCAAAAGCCCGGAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1764
 Db 1995 CGACACTGGGCCCCCTA-----CCGACGCTGGCTGCTGCTGCTGCTGCTGCTGCTG 2046
 QY 1765 ACCTGGCCCTGCTCAAGTGTCTGTCAGCAAGTCAAGCAGCAGCAGCAGCAGTGTAC 1824
 Db 2047 ACAGTGTCTTGTCCAGTGGCCGATCAGGCGCTATCCCGGCACTCCGCAATGACAG 2106
 QY 1825 CTGCAATCAACTGCTGCAAGTGTGATGATGATGATGATGATGATGATGATGATGATG 1884
 Db 2107 ATCCAGTCAACAGCTCGCTGATGACATCTCTGCTACAGGATGATCCCGGCGGAGC 2166
 QY 1885 AAGCCCTTGGTCAACAACCTGGTGTGAGGACTGGCTGCTGCTGCTGCTGCTGCTG 1944
 Db 2167 CGTCTGCTGCTGACGAGACCTGGGCTGAGGCGGCTGAGGCGGCTGAGGCGGCTG 2226
 QY 1945 ATGTGGATGACACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2004
 Db 2227 GTGTATGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2286
 QY 2005 TTCCAAAGGCTGATCCCGGAGTCCAGTCCATGACAGCAGCAGCAGCAGCAGCAGC 2064
 Db 2287 TCCACCGAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2346
 QY 2065 ATGATCTGCTATCGGGGCACTATGCTGCGGCAAGCAGCAGCAGCAGCAGCAGCAG 2124
 Db 2347 ATGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2406
 QY 2125 ATGGTGGCTTAAAGTGTGCAACAGGAGGCGGCGGAG-----CAAGATGCGAGCGG 2178
 Db 2407 CTAATGCGCTTAAAGTGTGCAACAGGAGGCGGCGGAG-----CAAGATGCGAGCGG 2466
 QY 2179 GTGAGCAATGTGACTCCAGAGCAACAGGCGGCGGCGGAGCAGCAGCAGCAGCAGC 2238
 Db 2467 GTTAGAGAGCTTGTCTCCAGAGCAACAGGCGGCGGCGGAGCAGCAGCAGCAGCAG 2526
 QY 2239 GCGGGGCGGCGGCGGCGGAG 2257

Db 2527 CCGGCCCCGAGCCCGCGG 2545
 RESULT 7
 US-10-245-851-99
 ; Sequence 99, Application US/10245851
 ; Publication No. US20030068782A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Baker, Kevin
 ; APPLICANT: Eaton, Dan
 ; APPLICANT: Filvaroff, Ellen
 ; APPLICANT: Goddard, Audrey
 ; APPLICANT: Grimaldi, J. Christopher
 ; APPLICANT: Gurney, Austin
 ; APPLICANT: Smith, Victoria
 ; APPLICANT: Stephan, Jean-Phillippe
 ; APPLICANT: Watanabe, Colin
 ; APPLICANT: Wood, William
 ; APPLICANT: Zhang, Zemin
 ; APPLICANT: Fong, Sherman
 ; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
 ; FILE OF INVENTION: ACIDS ENCODING THE SAME
 ; FILE REFERENCE: P3630R1C93
 ; CURRENT APPLICATION NUMBER: US/10/245,851
 ; PRIOR FILING DATE: 2002-09-16
 ; PRIOR APPLICATION NUMBER: 10/197942
 ; PRIOR FILING DATE: 2002-07-18
 ; PRIOR APPLICATION NUMBER: 60/059114
 ; PRIOR FILING DATE: 1997-09-17
 ; PRIOR APPLICATION NUMBER: 60/063046
 ; PRIOR FILING DATE: 1997-10-24
 ; PRIOR APPLICATION NUMBER: 60/065027
 ; PRIOR FILING DATE: 1997-11-10
 ; PRIOR APPLICATION NUMBER: 60/079689
 ; PRIOR FILING DATE: 1998-03-27
 ; PRIOR APPLICATION NUMBER: 60/086478
 ; PRIOR FILING DATE: 1998-05-22
 ; PRIOR APPLICATION NUMBER: 60/087607
 ; PRIOR FILING DATE: 1998-06-02
 ; PRIOR APPLICATION NUMBER: 60/089801
 ; PRIOR FILING DATE: 1998-06-18
 ; PRIOR APPLICATION NUMBER: 60/090557
 ; PRIOR FILING DATE: 1998-06-24
 ; PRIOR APPLICATION NUMBER: 60/090689
 ; PRIOR FILING DATE: 1998-06-25
 ; Remaining Prior Application data removed - See File Wrapper or PALM.
 ; NUMBER OF SEQ ID NOS: 116
 ; SEQ ID NO 99
 ; LENGTH: 2855
 ; TYPE: DNA
 ; ORGANISM: Homo Sapien
 ; US-10-245-851-99
 Query Match 19.8%; Score 621.8; DB 9; Length 2855;
 Best Local Similarity 61.3%; Pred. No. 1,3e-162;
 Matches 1078; Conservative 0; Mismatches 662; Indels 19; Gaps 4;

QY 507 GTTTGGCCGTTGCGAGCCCTGCCCCAAGTACTGTGCTGCCAGAAATCTGCTGAGTCACT 566
 Db 798 GTCAGCCACACCCAGGCCCATGTCCCCCGCTGCCCTGCCAGACACAGTCCGCTCCCT 857
 QY 567 GGGGACCTGTGCCCTCCAAAGGGGCTGCTTTTGTACCCCTGATATTGACCGGGGAC 626
 Db 858 AAGCGTGTGTGCCAGGGCAGCCCTCTGTTGTCGTCACCCCTGCTGAGCCCGCGG 917
 QY 627 AGTGGAGCTCGGCTGGGGCAACTTCATCATCCACATCAGCCCGCAGACTTTGCCAA 686
 Db 918 AGCCGAGCTCGGCTGCCAGCAACTTCATCGCTCCGTCGCGCGGCGGCGGCGG 977
 QY 687 CATCAGCGGCTGTGGACCTGACCCCTGTCCAGGAACACCATCAGCCATCCAGCCCTT 746
 Db 978 CATGACAGCCCTGTGTCATCTGAGCCTGTGCGGAAACACCATCCGCGCCGCTGCGG 1037

747 TTCCFTTGGACCTCGAGCCCTCCGCTCCCTGTCATTTGACAGCAATFCCGCTCCCAAG 806
 1038 CGCCTTCGCCACCTCGCGGCGCTCGTGCCTGCACCTGGATGCAACCGGCTGACCTC 1097
 807 CCTTGGGAGGACACCCCTCGGCGGCTGGTCAACCTGAGCACTTATGTTGAAACAA 866
 1098 ACTGGGCGAGGCGCAGCTGCGGGGCTGGTCAACTTTGGCCACCTCATCTCAGCAAA 1157
 867 CCAGCTGGGCGGCGCATCGAGATGAGGCTTTTGGAGACTTCTGTGACATGGAGATCT 926
 1158 CCAGCTGGCAGGGCTGGGGCGGCGCCTGGATGATTTGTCCGAGACACTGGAGACT 1217
 927 GGACTCTCCTAACAACTTCCATGGCTGCGGCTGGGACTCCGCTGGCAAGCATGGTCAA 986
 1218 CGACTCTCTTACAACTCGAGCTGCGGCTGGGACTGCACTTCCAACTCCAACTCCGCA 1277
 987 CCTCCAGCTGAGCTGGACCAACCTGCTGATCATCTGCGGAGGCGACTTTGC 1046
 1278 CGTCAACAGTTGGGCGCTCGAACAACTGCTGGCTTCTGTGCCCC--GGCGCTTTTC 1334
 1047 AGACTGAGAACTGGCGCGCTGGATCTCACTCCAACTCGGCTGCAGAGCTGCCCCC 1106
 1335 CGGCTTGCACAGCTGGCCGCTGGACATGACTTCCAACTCCAACTCCAACTCCGCA 1394
 1107 TGATCCCATTTTGGCGGCTCCAGGCTTCCGCTTTTGAAGCACCACCCCTTTGCCCCACC 1166
 1395 CGACCCTCTTCTCCCGCTGCTCCCTGCTGCTCAGGCGCCCGGGCTGCGCCGCTTGC 1454
 1167 CTGTGCTTTAGTTTTGGGGTAAACCACTTCACTGCAATTTGTAGCTTCTTGTGCTGG 1226
 1455 CCGTGTGTGGCTTTTGGCGGAAACCCCTGCACTGCACTTCCGAGCTGGTGTGGCTGG 1514
 1227 GAGGCTCGAGGGGACGACTGCTGAAACCTGCTGCTCCCGAGGGGCTCAGGGTGC 1286
 1515 TCGCTTGGCGGGAGGAGCACTCGAGGCTTCCGCGTCCCGCCTGCTTGGGGGGCG 1574
 1287 CTACTTCTGGCATGTGCGTAGGAGGTTGTTGCGAGCGGCTTCTCATCACCAGCA 1346
 1575 CTACTTCTGGGGGTGGCGAGGAGTTTGTCTCGAGCGCGCGGCTGCTGACTCACCG 1634
 1347 CACACAAATGCTGTTGCGGCGGCGGCGGCGCAGCTCAAGTGCAAGGCAATGG 1406
 1635 CTCACCACTGCTGGCTGTCGCGGAGCTGGCGGCTGCTGCTGCTGCTGCGGCGAGTGG 1694
 1407 GGACCCAGCCCTTATCCACTGGTAGCCCGGATGAGCCCGCTGGTAGGAACTCCTC 1466
 1695 GGACCCAGGCGGCTGCTGGTTGGTGTACCCAGGCGCGGCTGTAGGCAACTCAAG 1754
 1467 AAGGACCGCTGTATGCAATGGCACCCTGGACATTTTCAACCACATCTCAGGACAG 1526
 1755 CGGTGCGCGCGCTTCCCAATGGGACGCTGGAGCTGCTGGTCAACCAGCGGTGATGG 1814
 1527 TGGTCCCTCACCTGATGCTCCAAATGCTGCGGAGGCGGCGGCGGAGGCTGGAGGT 1586
 1815 TGGCATTTTCCACTGATTTGGCGGCAATGCACTGGCGAGGCGCACAGTGTGGAGCT 1874
 1587 CTCATCGTCCAGTGGCACACTCAGCAACAGCAGCCAGCCGCA--CTGACCCCGCAAG 1644
 1875 GACTGTGGTCCCAACACTCTCAGTAGCCAAAGCACGCTGTGACCCCGCGG 1934
 1645 TCCCGCTCTCAGACATCACTGGCTCCAGCAAGACGAGCGCGGGGAGGTGGAGGAGTGG 1704
 1935 GGACCGGGATCTGTGTGCTTCAACCCCTCGCTGCTGCTGCTGCTGCTGCCAAGTGGC 1994
 1705 GCGGAGAGCTTCCAAAGCCCCCGGAAAGCGGCTGTGCTGTTGTGAAGTGACACC 1764
 1995 CGACACTGGGCGCCCTA-----CCGACCGTGGCTCCAGGCTGACTGAGCAGCGGGG 2046
 1765 ACCTGGGCGCTTCAAGTGTCTGTGACCAAGTCCAGCACTCCCGGCTGAGGATGATCCAG 1824
 2047 ACAGTGTCTTGTTCAGTGGCGGATCAGCGGCTTATCCCGGCGATCCGCATGTACCAG 2106

1825 CTGAGTACAACTCTGAGCAGTGGAGTCTGATTTACAGGATGATCCAGCCTTCCAAC 1884
 2107 ATCCAGTACAACTCGGCTGACTGACATCCTCGTCTACAGGATGATCCCGCGGAGAGC 2166
 1885 AAGCCCTTGTGTCACAACTCGGCTGAGGACTGGCTTGTGTGTGCTGGCC 1944
 2167 CGCTGTTCTGTCAGGACTGCGGCTGAGGCTGAGGACTGAGTCTGTCGCTGCTGCC 2226
 1945 ATGTGGATGACAGCCAGCAGCACTCAGCGCCACCAACATCTGGGGCTGGCCCGATTC 2004
 2227 GTGTATGAGGACAGCGCCACGGGCTCAGCGCCAGCGGCTGTGGGCTGGCCCGCTTC 2286
 2005 TTACCAAGGCTGACTACCCGAGTGCAGTCCATGCAAGCCAGATTTCTGGGCGGCGCC 2064
 2287 TCCACCGAAGCTGCGGCTGCGGCCATGCGGGGCGCGCACGCTCCCTTCTGGGCGGCG 2346
 2065 ATGATCTGTGTCATCGGGGCTCATCGTGCAGGCTGCTGCTTATCGTCAATCCTC 2124
 2347 ATGATCATCGGCTGGGGGCTCATCTAGCTCGGTACTGTTCTTTCATCTTCTGCTG 2406
 2125 ATGTGCGCTACAAGTCTGCAACCAAGGCGGCGGCTGCGGCTTCCGAGCGCGG 2178
 2407 CTAATGCGCTACAAGTGCACGGGCGGCGGCGGCGGCAAGGCTTCCCGGCGCT 2466
 2179 GTGAGCAATGTTACTCCAGCAACCGGCGGCGGCGGCGGCGGCTTCCAGAGCGGCGCA 2238
 2467 GTTAGCAGGCTTGTCTCCAGCAACCGGCGGCGGCGGCGGCGGCGGCGGCGG 2526
 2239 CCGGGGCGGCGGCGGCGGCGG 2257
 2527 CCGGCGGCGGCGGCGGCGG 2545

RESULT 8

US-10-245-883-99
 ; Sequence 99, Application US/10245883
 ; Publication No. US20030068783A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Baker, Kevin
 ; APPLICANT: Eaton, Dan
 ; APPLICANT: Filvaroff, Ellen
 ; APPLICANT: Goddard, Audrey
 ; APPLICANT: Grimaldi, J. Christopher
 ; APPLICANT: Gurney, Austin
 ; APPLICANT: Smith, Victoria
 ; APPLICANT: Stephan, Jean-Phillippe
 ; APPLICANT: Watanabe, Colin
 ; APPLICANT: Wood, William
 ; APPLICANT: Zhang, Zemin
 ; APPLICANT: Fong, Sherman
 ; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
 ; FILE REFERENCE: P3630R1C70
 ; CURRENT APPLICATION NUMBER: US/10/245,883
 ; CURRENT FILING DATE: 2002-09-16
 ; PRIORITY APPLICATION NUMBER: 10/197942
 ; PRIORITY FILING DATE: 2002-07-18
 ; PRIORITY APPLICATION NUMBER: 60/059114
 ; PRIORITY FILING DATE: 1997-09-17
 ; PRIORITY APPLICATION NUMBER: 60/063046
 ; PRIORITY FILING DATE: 1997-10-24
 ; PRIORITY APPLICATION NUMBER: 60/065027
 ; PRIORITY FILING DATE: 1997-11-10
 ; PRIORITY APPLICATION NUMBER: 60/079689
 ; PRIORITY FILING DATE: 1998-03-27
 ; PRIORITY APPLICATION NUMBER: 60/086478
 ; PRIORITY FILING DATE: 1998-05-22
 ; PRIORITY APPLICATION NUMBER: 60/087607
 ; PRIORITY FILING DATE: 1998-06-02
 ; PRIORITY APPLICATION NUMBER: 60/089801
 ; PRIORITY FILING DATE: 1998-06-18
 ; PRIORITY APPLICATION NUMBER: 60/090557
 ; PRIORITY FILING DATE: 1998-06-24

; PRIOR APPLICATION NUMBER: 60/090689
 ; PRIOR FILING DATE: 1998-06-25
 ; Remaining Prior Application data removed - See File Wrapper or PALM.
 ; NUMBER OF SEQ ID NOS: 116
 ; SEQ ID NO 99
 ; LENGTH: 2855
 ; TYPE: DNA
 ; ORGANISM: Homo Sapien
 ; US-10-245-883-99

Query Match 19.8%; Score 621.8; DB 9; Length 2855;
 Best Local Similarity 61.3%; Pred. No. 1.3e-162;
 Matches 1078; Conservative 0; Mismatches 662; Indels 19; Gaps 4;

QY 507 GTTTCGGTGTGAGCGCTGCCCAAGTACTGTCTGCCAGAAATCTGTCTGAGTCACT 566
 Db 798 GTACGCCACACCCAGCCCAATGTCCTCCCGCGCTGCGCTGCCAGACACAGTCCGCCCT 857
 QY 567 GGGACCCCTGTGCCCTCCAAAGGGGCTGCTTTGFPACCCCTGATATTGACCCGGGAC 626
 Db 858 AAGCTGTCTGTCCAGGGGAGGCTTCTGTTGTCGTCACCCCTCGCTGGACCCGGCG 917
 QY 627 AGTGGAGCTGGCGGGAATCTATCATCCATCAGCGCCAGGACTTTGCCAA 686
 Db 918 AGCCGAGCTGGCTGACAGACAATTCATGCTCCGTCGCGCCGCGGACTGGCCAA 977
 QY 687 CATGACGGGCTGTGGACCTGACCTCTCCAGGAACACCATCAGCCATCCAGCCCTT 746
 Db 978 CATGACGAGCTGTGATCTGAGCTGTGGGGAACACCATCCGCGCAGCTGGCTGCCGG 1037
 QY 747 TTCCTTTCTGGACCTGAGAGCTCCGCTCCCTGTCATCTGACGCAATCCAGCCCTT 746
 Db 1038 CGCCTTCCCGGCTGCGGGCCCTGCGTGCCTGCACTGGGCAACCTCATCTCAGCAACA 1157
 QY 807 CTTGGGAGACACCTCCGGGCTGTGTCACCTGACCTGACCTTATCGTAAACA 866
 Db 1098 ACTGGGAGGCGAGCTGGCGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1217
 QY 867 CCAGCTGGGGGCGATCGAGATAGGCTTTTGGAGCTTCTGCTGACATTTGAGGATCT 926
 Db 1158 CCAGCTGGAGGCTGGCGGCGGCGGCTGGATGATTTGTCGCGAGACACTGGAGACCT 1217
 QY 927 GGACCTTCTTACAAACCTCCATGGCTCCGCTGGGACTCCGTCGGGAGGCGCATGCTCAA 986
 Db 1218 CGACTCTTCTTACAAACCTCCATGGCTCCGCTGGGACTCCGTCGGGAGGCGCATGCTCAA 1277
 QY 987 CCTCCACAGCTGAGCTGGACCAACCTGCTGATCACAATCGCCAGGCGACCTTTGC 1046
 Db 1278 CGTCAACAGCTGGGCTCGACCAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1334
 QY 1047 AGACTGCAAACTGGCCCGCTGGATCTCACCTCCAATCGGCTGCAGAGCTGCCCCC 1106
 Db 1335 CGGCTGCAAACTGGCCCGCTGGATCTCACCTCCAATCGGCTGCAGAGCTGCCCCC 1394
 QY 1107 TGATCCACTTTTGGCCGCTCCAGGCTTGGCTTTGACAGCCACACCTTTGCCCCACC 1166
 Db 1395 CGACCACTTCTTCCCGCTGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1454
 QY 1167 CTTGCTTTTGTGGGGTAAACCCTTCACTGCAATTTGAGCTTCTGCTGCTGCTGCTGCT 1226
 Db 1455 CTTGCTTTTGTGGGGTAAACCCTTCACTGCAATTTGAGCTTCTGCTGCTGCTGCTGCT 1514
 QY 1227 GAGGCTGAGCGGAGATGACTGGAAACCTGTGGCTCCCGGGGGGCTCAAGGGTGC 1286
 Db 1515 TCGCCCTGGCGGGAGAGACCTCGAGGCTTGGGCTCCCGGCTGCTGCTGCTGCTGCTGCT 1574
 QY 1287 CTACTTCTGGCATGCTGGAGGAGTTTGTGCGGAGCGGCTTCTCATCCACCAGCA 1346
 Db 1575 CTACTTCTGGCGGTTGGCGGAGGAGTTTGTGCGGAGCGGCTTCTCATCCACCAGCG 1634
 QY 1347 CACACAACTTCTGTTGAGGCGGAGGCGGACACTCAAGTGCAAAAGCAATGG 1406
 Db 1635 CTACCACTCTGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1694

1407 GGACCCAGCCCTTATCCACTGGGTAGCCCGGATGACCCGCTGTTGGTGGAACTCCTC 1466
 Db 1695 GGACCCAGAGCCCGTGTGGTGTCAACCCAGGCGGCTGTAGGCACTCAAG 1754
 QY 1467 AAGGACCGCTGTATGACAATGGACCCCTGGACATTTTCATCACCACATCTCAGGACAG 1526
 Db 1755 CCGTCCCGGCTTCCCAATGGGACGCTGAGTGTCCGACCCGAGGCTGGTGTATG 1814
 QY 1527 TGGTGCCTTCACTGCAATGCTGCAATGCTCCGAGAGGCGGAGGCGGATGGTGGAGGT 1586
 Db 1815 TGGCATTTCACTGCAATGCTGCAATGCTCCGAGAGGCGGAGGCGGATGGTGTGGAGCT 1874
 QY 1587 CTCATCGTCCAGCTGCCACACTCAGCAACAGCAGCAGCCGCA--CTCCACCCCCCAAG 1644
 Db 1875 GACTGTGGTCCCGCCACCTCTCAGCTAGCAACAGCAGCAGCTGTGACCCCGCG 1934
 QY 1645 TCCCGCTCTCAGACATCACTGGCTCAGCAAGACAGCAGGCGGAGGTGGAGCAGTGGG 1704
 Db 1935 GGACGGGATCTGATGCTTCAACCCCGCTCCGCTGCTTCTGCTTCTGCTGCTGCTGCTG 1994
 QY 1705 GCGGAGAGCTCCCAAAAGCCCGGAAAGCGGCTGTGCTTGTGCTGAAAGTACACC 1764
 Db 1995 CGACACTTGGGCCCCCTA-----CCGACCGTGGCTCCAGGTGACTGAGCAGCGGGCC 2046
 QY 1765 ACCTCGGCTTGGTCAAGTGTGTGCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 1824
 Db 2047 ACAGCTCTTGTTCAGTGGCGGATCAGCGGCTATCCCGGGCACTCCGATGTGCTGCTG 2106
 QY 1825 CTGAGTAACTGCTGTGACGATGAGTACTGATTTACAGGATGATCCAGGCTTCCAAC 1884
 Db 2107 ATCCAGTAACTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2166
 QY 1885 AAGCCTTCTGTGTCAAAACCTGGTGTGAGGACTGGTACGACTGTGCTGCTGCTGCTG 1944
 Db 2167 CGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2226
 QY 1945 ATGTGGATGACAGCCAGCAGACACTCAGGCGCAGCAACATCTGTGGGCTGGCCAGTTC 2004
 Db 2227 GTGTATGAGCAGCGCCAGCGGGGCTCAGGCGCAGCGGGGCTGTGGGCTGTGGGCTG 2286
 QY 2005 TTCACCAAGGCTGACTACCCGAGTGGCCAGTCCATGACAGCCAGATTTCTGGGGGAC 2064
 Db 2287 TCCACCGA--CTGCTGGCTGGCCATGCGGGGCGCGCAGCTCCCTTCTGGGCGGACG 2346
 QY 2065 ATGATCTTGGTATCGGGGCAATCATCTGCTGCGGCTGCTGCTGCTGCTGCTGCTGCT 2124
 Db 2347 ATGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2406
 QY 2125 ATGGTGGCTTCAAGGTCTGCAACCAAGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2178
 Db 2407 CTAATGGGCTCAAGGTGCAAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2466
 QY 2179 GTGAGCAATGTTGCTGCGAGACCAAGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2238
 Db 2467 GTTAGCAGGTTTGTCTCCAGACCAAGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2526
 QY 2239 GCGGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2257
 Db 2527 CCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2545

RESULT 9
 US-10-237-535-99
 ; Sequence 99, Application US/10237535
 ; Publication No. US20030073188A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Baker, Kevin
 ; APPLICANT: Eaton, Dan
 ; APPLICANT: Filvaroff, Ellen
 ; APPLICANT: Goddard, Audrey
 ; APPLICANT: Grimaldi, J. Christopher
 ; APPLICANT: Gurney, Austin

```

APPLICANT: Smith,Victoria
APPLICANT: Stephan,Jean-Phillippe
APPLICANT: Watanabe,Colin
APPLICANT: Wood,William
APPLICANT: Zhang,Zemin
APPLICANT: Fong,Sherman
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3630R1C3
CURRENT APPLICATION NUMBER: US/10/237,535
CURRENT FILING DATE: 2002-09-06
PRIOR APPLICATION NUMBER: 10/197942
PRIOR FILING DATE: 2002-07-18
PRIOR APPLICATION NUMBER: 60/059114
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/063046
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/065027
PRIOR FILING DATE: 1997-11-10
PRIOR APPLICATION NUMBER: 60/079689
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/086478
PRIOR FILING DATE: 1998-05-22
PRIOR APPLICATION NUMBER: 60/087607
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/089801
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/090557
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090689
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/091358
PRIOR FILING DATE: 1998-07-01
PRIOR APPLICATION NUMBER: 60/091978
PRIOR FILING DATE: 1998-07-07
PRIOR APPLICATION NUMBER: 60/099803
PRIOR FILING DATE: 1998-09-10
PRIOR APPLICATION NUMBER: 60/106932
PRIOR FILING DATE: 1998-11-03
PRIOR APPLICATION NUMBER: 60/115554
PRIOR FILING DATE: 1999-01-12
PRIOR APPLICATION NUMBER: 60/119342
PRIOR FILING DATE: 1999-02-09
PRIOR APPLICATION NUMBER: 60/123957
PRIOR FILING DATE: 1999-03-12
PRIOR APPLICATION NUMBER: 60/123972
PRIOR FILING DATE: 1999-03-11
PRIOR APPLICATION NUMBER: 60/127372
PRIOR FILING DATE: 1999-04-01
PRIOR APPLICATION NUMBER: 60/131271
PRIOR FILING DATE: 1999-04-27
PRIOR APPLICATION NUMBER: 60/133459
PRIOR FILING DATE: 1999-05-11
PRIOR APPLICATION NUMBER: 60/135725
PRIOR FILING DATE: 1999-05-25
PRIOR APPLICATION NUMBER: 60/135729
PRIOR FILING DATE: 1999-05-25
PRIOR APPLICATION NUMBER: 60/135750
PRIOR FILING DATE: 1999-05-25
PRIOR APPLICATION NUMBER: 60/138385
PRIOR FILING DATE: 1999-06-09
PRIOR APPLICATION NUMBER: 60/140653
PRIOR FILING DATE: 1999-06-22
PRIOR APPLICATION NUMBER: 60/141037
PRIOR FILING DATE: 1999-06-23
PRIOR APPLICATION NUMBER: 60/144732
PRIOR FILING DATE: 1999-07-20
PRIOR APPLICATION NUMBER: 60/144758
PRIOR FILING DATE: 1999-07-20
PRIOR APPLICATION NUMBER: 60/144790
PRIOR FILING DATE: 1999-07-20
PRIOR APPLICATION NUMBER: 60/145228
PRIOR FILING DATE: 1999-07-20
PRIOR APPLICATION NUMBER: 60/145698
PRIOR FILING DATE: 1999-07-26
PRIOR APPLICATION NUMBER: 60/146222
PRIOR FILING DATE: 1999-07-28
PRIOR APPLICATION NUMBER: 60/146843
PRIOR FILING DATE: 1999-08-03
PRIOR APPLICATION NUMBER: 60/148188
PRIOR FILING DATE: 1999-08-10
PRIOR APPLICATION NUMBER: 60/148513
PRIOR FILING DATE: 1999-08-12
PRIOR APPLICATION NUMBER: 60/149327
PRIOR FILING DATE: 1999-08-17
PRIOR APPLICATION NUMBER: 60/149395
PRIOR FILING DATE: 1999-08-17
PRIOR APPLICATION NUMBER: 60/150114
PRIOR FILING DATE: 1999-08-20
PRIOR APPLICATION NUMBER: 60/151700
PRIOR FILING DATE: 1999-08-31
PRIOR APPLICATION NUMBER: 60/151734
PRIOR FILING DATE: 1999-08-31
PRIOR APPLICATION NUMBER: 60/162506
PRIOR FILING DATE: 1999-10-29
PRIOR APPLICATION NUMBER: 60/170262
PRIOR FILING DATE: 1999-12-09
PRIOR APPLICATION NUMBER: 60/177118
PRIOR FILING DATE: 2000-01-20
PRIOR APPLICATION NUMBER: 60/179851
PRIOR FILING DATE: 2000-02-02
PRIOR APPLICATION NUMBER: 60/180921
PRIOR FILING DATE: 2000-02-08
PRIOR APPLICATION NUMBER: 60/187202
PRIOR FILING DATE: 2000-03-03
PRIOR APPLICATION NUMBER: 60/198587
PRIOR FILING DATE: 2000-04-18
PRIOR APPLICATION NUMBER: 60/199614
PRIOR FILING DATE: 2000-04-25
PRIOR APPLICATION NUMBER: 60/206330
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/206368
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/209832
PRIOR FILING DATE: 2000-06-05
PRIOR APPLICATION NUMBER: 60/218371
PRIOR FILING DATE: 2000-07-13
PRIOR APPLICATION NUMBER: 60/223695
PRIOR FILING DATE: 2000-08-02
PRIOR APPLICATION NUMBER: 60/229896
PRIOR FILING DATE: 2000-09-01
PRIOR APPLICATION NUMBER: 60/230621
PRIOR FILING DATE: 2000-09-05
PRIOR APPLICATION NUMBER: 60/232887
PRIOR FILING DATE: 2000-09-15
PRIOR APPLICATION NUMBER: 60/235147
PRIOR FILING DATE: 2000-09-22
PRIOR APPLICATION NUMBER: 60/261878
PRIOR FILING DATE: 2001-01-12
PRIOR APPLICATION NUMBER: 60/261910
PRIOR FILING DATE: 2001-01-16
PRIOR APPLICATION NUMBER: 60/261939
PRIOR FILING DATE: 2001-01-16
PRIOR APPLICATION NUMBER: 60/262150
PRIOR FILING DATE: 2001-01-16
PRIOR APPLICATION NUMBER: 60/264395
PRIOR FILING DATE: 2001-01-25
PRIOR APPLICATION NUMBER: 60/266421
PRIOR FILING DATE: 2001-02-02
PRIOR APPLICATION NUMBER: 60/267623
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/274399
PRIOR FILING DATE: 2001-03-09
PRIOR APPLICATION NUMBER: 60/280982
PRIOR FILING DATE: 2001-04-03
PRIOR APPLICATION NUMBER: 60/282129

```

867 CCAGCTGGGGGCGATCGCAGATGAGGCTTTTGGAGGATTCCTCTGCTGACATTTGGAGATCT 926
 1158 CCAGCTGGGAGGGCTGGCGCGCGCCCTGATGATGATGTCGGAGACACTGGAGGACCT 1217
 927 GGACCTCTCTTCAACAACAACCTCCATGGCTGGCGTGGGACTCCGTCGCGAGCATGATGTCAA 986
 1218 CGACCTCTCTTCAACAACAACCTCCGAGCAGCTGGCTGGGAGGCTGGGCGCTGGGCAA 1277
 987 CTTCCACAGCTGAGCCTGGACCAACCTGCTGGATCATCGCCGAGGCGACCTTTTCG 1046
 1278 CGTCAACACGTTGGGCTCGACCAACCTGCTGGCTTCTGTGCC---GGCGTTTTTC 1334
 1047 AGACCTGCAGAAACTGGCGCGCTGGATCTCACTCCATTCGGCTGCAGAAAGTGCAGCC 1106
 1335 CCGCTGCACAGCTGGCGCGCTGGACATGACCTCAACCGCTGACCAAAATCCACC 1394
 1107 TGATCCCATCTTTGCCCGCTCCAGGCTTCCGCTTTCAGAGCCACACCTTTGCCCCACC 1166
 1395 CGACCACTCTTCCCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCG 1454
 1167 CTTGCTCTTGTGGGGTAAACCACTTCACTGCAATTTGAGCTTCTCTGGTGGC 1226
 1455 CCTGGTGTGCTGGCTTGGCGGAAACCCCTTGCATGCACTGCAACTGCGAGCTGGTGGTGG 1514
 1227 GAGGCTCGAGCGGACGATGACCTTGGAAACCTGTGGCTTCCCGAGGGGCTCAAGGCTG 1286
 1515 TCGCTGGCGGGGAGGACGACCTCGAGGCTGGCTTCCGCTTCCGCTTCCGCTTCCGCT 1574
 1287 CTACTTCTGGCATGTGGTGGAGGAGTTGTGGAGGCTGGCTTCCGCTTCCGCTTCCGCT 1634
 1575 CTACTTCTGGCGGCTGGCGGAGGAGTTGTGGAGGCTGGCTTCCGCTTCCGCTTCCGCT 1694
 1347 CACACAAAGTTGCTGGTTCTGGAGGCGGAGGCGGCTTCCGCTTCCGCTTCCGCTTCCG 1406
 1635 CTCACCACTTGGCTTGGCTGGCGGAGGCTGGCTTCCGCTTCCGCTTCCGCTTCCGCT 1694
 1407 GGACCCAGCCCTTATCCACTGGGTAGCCCCCGATGACCGCTGGTGGTGGTGGTGGTGG 1466
 1695 GGACCCAGGCGGCTGGCTGGCTGGCGGAGGCTGGCTTCCGCTTCCGCTTCCGCTTCCG 1754
 1467 AAGGACCGCTGCTATGACAAATGGACCTGGACATCTTTCATCACCACTCTCAGGACAG 1526
 1755 CCGTGGCGGCGCTTCCCAATGGACGCTGGAGCTGGTGGTGGTGGTGGTGGTGGTGG 1814
 1527 TGGTGCCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1586
 1815 TGGCATCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1874
 1587 CTCATCGTCCAGCTGCCACACCTCAGCAACAGCAGCAGCCGCA--CTGCACCCCGGCAAG 1644
 1875 GACTGTGGGTCCCGCCACCACTTCTCAGTACGCAACAGCAGCAGCTGTGACCCCGCG 1934
 1645 TCCCGCTCTCAGACATCACTGGCTCCAGCAAGACAGCCGCGGGAGGTGGAGGAGTGG 1704
 1935 GGACGGGATCTGATGCTTCCACCCCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCT 1994
 1705 GCGGAGAGCTCCCAAAAGCCCCCGGNAACGGGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1764
 1995 CGACACTGGGCGGCTT-----CCGACCGTGGCTTCCGCTTCCGCTTCCGCTTCCGCT 2046
 1765 ACCTCGGCTGGTCAAGTGTGTGTCAGAAAGTTCAGACACCCCGGCTGAAGTGTACACAG 1824
 2047 ACAGCTGCTTGTCCAGTGGCGGATCAGCGGCTTCCGCGGCTTCCGCGGCTTCCGCGG 2106
 1825 CTCAGTACAACTGCTCTCAGATGAGTACTGATTTAAGGATGATCCAGGCTTCCAGC 1884
 2107 ATCCAGTACAACTGCTCTCAGATGAGTACTGATTTAAGGATGATCCAGGCTTCCAGC 2166
 1885 AAGGCTTGTGTGTCACAACTGCTGTCAGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1944
 2167 CGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2226

Qy 867
 Db 1158
 Qy 927
 Db 1218
 Qy 987
 Db 1278
 Qy 1047
 Db 1335
 Qy 1107
 Db 1395
 Qy 1167
 Db 1455
 Qy 1227
 Db 1515
 Qy 1287
 Db 1575
 Qy 1347
 Db 1635
 Qy 1407
 Db 1695
 Qy 1467
 Db 1755
 Qy 1527
 Db 1815
 Qy 1587
 Db 1875
 Qy 1645
 Db 1935
 Qy 1705
 Db 1995
 Qy 1765
 Db 2047
 Qy 1825
 Db 2107
 Qy 1885
 Db 2167

Qy 507
 Db 798
 Qy 567
 Db 858
 Qy 627
 Db 918
 Qy 687
 Db 978
 Qy 747
 Db 1038
 Qy 807
 Db 1098

Prior Filing Date: 2001-04-04
 Prior Application Number: 60/282199
 Prior Filing Date: 2001-04-04
 Prior Application Number: 60/290589
 Prior Filing Date: 2001-05-09
 Prior Application Number: 09/180997
 Prior Filing Date: 1998-11-19
 Prior Application Number: 09/267213
 Prior Filing Date: 1999-03-12
 Prior Application Number: 09/380137
 Prior Filing Date: 1999-08-25
 Prior Application Number: 09/380138
 Prior Filing Date: 1999-08-25
 Prior Application Number: 09/403297
 Prior Filing Date: 1999-10-18
 Prior Application Number: 09/423741
 Prior Filing Date: 1999-11-10
 Prior Application Number: 09/709238
 Prior Filing Date: 2000-11-08
 Prior Application Number: 09/802706
 Prior Filing Date: 2001-03-09
 Prior Application Number: 09/872035
 Prior Filing Date: 2001-06-01
 Prior Application Number: 09/918585
 Prior Filing Date: 2001-07-30
 Prior Application Number: 09/924419
 Prior Filing Date: 2001-08-06
 Prior Application Number: 09/927796
 Prior Filing Date: 2001-08-09
 Prior Application Number: 09/929404
 Prior Filing Date: 2001-08-13
 Prior Application Number: 09/931836
 Prior Filing Date: 2001-08-16
 Prior Application Number: 09/941992
 Prior Filing Date: 2001-08-28
 Prior Application Number: 09/946374
 Prior Filing Date: 2001-09-04
 Prior Application Number: 10/001054
 Prior Filing Date: 2001-11-30
 Prior Application Number: 10/052586
 Prior Filing Date: 2002-01-15
 Prior Application Number: 10/081056
 Prior Filing Date: 2002-02-20
 Prior Application Number: 10/119480
 Prior Filing Date: 2002-04-09

Query Match 19.8%; Score 621.8; DB 9; Length 2855;
 Best Local Similarity 61.3%; Pred. No. 1.3e-162;
 Matches 1078; Conservative 0; Mismatches 662; Indels 19; Gaps 4;


```

Qy 1945 ATGCGGATGACACAGCCACGACACTACCGCCACCAACATCGTGGGCTGCAGCTTC 2004
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2227 GTGTATGAGGACAGCCGCGCGCTACCGGCTCAAGCGGCTGTGGGCTGCAGCTTC 2286
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 2005 TTCACCAAGCTGACTACCCGAGTGCCAGTCCATGACAGCAGAGTTCGCGGCACC 2064
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2287 TCCACCGAACCTGCGCTGGGCGGCTGGGCGCCGCGCAGCTCCCTTCCTGGCGGCAG 2346
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 2065 ATGATCTGTGTCATCGGGGGCATCATCGTGGCGCGCTGCTGTTTCATCGTATCCTC 2124
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2347 ATGATCATCGCGCTGGCGGCGTCATCGTAGCCTCGGTACTGTTCTTCATCTTCGTGCTG 2406
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 2125 ATGGTGGCTACAAAGTCTGCAACACACAGGCCCCAG-----CAAGATGCCAGCGCC 2178
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2407 CTAATGCGCTCAAGTGCACGGCGGCGCGCGCCAGCCCCCGCGCAAGATTCGCCGCGCT 2466
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 2179 GTGACCAATGTCTACTCGCAGAACCAAGCGCGCCAGCCCGCTCCCAAGCAGCCACCA 2238
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2467 GTTACAGGTTTGTCTCCAGAACCAAGCGCGCTGGCGCCCGCGCCAGCCCGCCCGG 2526
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 2239 GCCGGGCCCCCGCGCG 2257
    ||| ||| ||| |||
Db 2527 CCGGCCCCGGAGCCCGCGG 2545
    ||| ||| ||| |||

RESULT 10
US-10-238-183-99
; Sequence 99, Application US/10238183
; Publication No. US20030073189A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin
; APPLICANT: Eaton, Dan
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Goddard, Audrey
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Phillippe
; APPLICANT: Watambe, Colin
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; APPLICANT: Fong, Sherman
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3630RIC11
; CURRENT APPLICATION NUMBER: US/10/238,183
; CURRENT FILING DATE: 2002-09-09
; PRIOR APPLICATION NUMBER: 10/197942
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/059114
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/063046
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/065027
; PRIOR FILING DATE: 1997-11-10
; PRIOR APPLICATION NUMBER: 60/079689
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/086478
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090689
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/091358
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091978
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/099803
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: 60/106932
; PRIOR FILING DATE: 1998-11-03
; PRIOR APPLICATION NUMBER: 60/115554
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: 60/119342
; PRIOR FILING DATE: 1999-02-09
; PRIOR APPLICATION NUMBER: 60/123957
; PRIOR FILING DATE: 1999-03-12
; PRIOR APPLICATION NUMBER: 60/123972
; PRIOR FILING DATE: 1999-03-11
; PRIOR APPLICATION NUMBER: 60/127372
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/131271
; PRIOR FILING DATE: 1999-04-27
; PRIOR APPLICATION NUMBER: 60/133459
; PRIOR FILING DATE: 1999-05-11
; PRIOR APPLICATION NUMBER: 60/135725
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: 60/135729
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: 60/135750
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: 60/138385
; PRIOR FILING DATE: 1999-06-09
; PRIOR APPLICATION NUMBER: 60/140653
; PRIOR FILING DATE: 1999-06-22
; PRIOR APPLICATION NUMBER: 60/141037
; PRIOR FILING DATE: 1999-06-23
; PRIOR APPLICATION NUMBER: 60/144732
; PRIOR FILING DATE: 1999-07-20
; PRIOR APPLICATION NUMBER: 60/144758
; PRIOR FILING DATE: 1999-07-20
; PRIOR APPLICATION NUMBER: 60/144790
; PRIOR FILING DATE: 1999-07-20
; PRIOR APPLICATION NUMBER: 60/145228
; PRIOR FILING DATE: 1999-07-20
; PRIOR APPLICATION NUMBER: 60/145698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: 60/146222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: 60/146843
; PRIOR FILING DATE: 1999-08-03
; PRIOR APPLICATION NUMBER: 60/148188
; PRIOR FILING DATE: 1999-08-10
; PRIOR APPLICATION NUMBER: 60/148513
; PRIOR FILING DATE: 1999-08-12
; PRIOR APPLICATION NUMBER: 60/149327
; PRIOR FILING DATE: 1999-08-17
; PRIOR APPLICATION NUMBER: 60/149395
; PRIOR FILING DATE: 1999-08-17
; PRIOR APPLICATION NUMBER: 60/150114
; PRIOR FILING DATE: 1999-08-20
; PRIOR APPLICATION NUMBER: 60/151700
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: 60/151734
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: 60/162506
; PRIOR FILING DATE: 1999-10-29
; PRIOR APPLICATION NUMBER: 60/170262
; PRIOR FILING DATE: 1999-12-09
; PRIOR APPLICATION NUMBER: 60/177118
; PRIOR FILING DATE: 2000-01-20
; PRIOR APPLICATION NUMBER: 60/179851
; PRIOR FILING DATE: 2000-02-02
; PRIOR APPLICATION NUMBER: 60/180921
; PRIOR FILING DATE: 2000-02-08
; PRIOR APPLICATION NUMBER: 60/187202
; PRIOR FILING DATE: 2000-03-03
; PRIOR APPLICATION NUMBER: 60/198587
; PRIOR FILING DATE: 2000-04-18
; PRIOR APPLICATION NUMBER: 60/199614
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 60/206330

```

PRIOR APPLICATION NUMBER: 10/001054
 PRIOR FILING DATE: 2001-11-30
 PRIOR APPLICATION NUMBER: 10/052586
 PRIOR FILING DATE: 2002-01-15
 PRIOR APPLICATION NUMBER: 10/081056
 PRIOR FILING DATE: 2002-02-20
 PRIOR APPLICATION NUMBER: 10/119480
 PRIOR FILING DATE: 2002-04-09

Query Match 19.8%; Score 621.8; DB 9; Length 2855;
 Best Local Similarity 61.3%; Pred. No. 1.3e-162;
 Matches 1078; Conservative 0; Mismatches 662; Indels 19; Gaps 4;

507 GTTTGGCGGTGTCGACGCTGCCCCAAAGTACTGTGTCTCCAGAAATCTCTCTGAGTCACT 566
 798 GTCAGCCACACCCAGCCCATGTCCCCCGCGCTCCAGACACAGTGGCTGCCCCCT 857
 567 GGGGACCCCTGTGCCCTCCAAAGGGGTGCTTTTGTACCCCTGATATTGACCGGGGAC 626
 858 AAGCGTGTGTGCCAGGGGAGGCTCCCTGTTGTCGCCACCTCGCTGGACCGCGGGC 917
 627 AGTGGAGCTCGGCTGGCGGCAACTTTCATFCCACATCAGCCGCGCAGACTTTGCCAA 686
 918 AGCGAGCTGGCTGGCAGCAACTTTCATCGCTCCGTCGCGCCGCGCGGCACTGGCCAA 977
 687 CATGACGGGCTGTGGACCTGACCTGTCAGAAACACCATCAGCCACATCCAGCCCTT 746
 978 CATGACAGGCTGTGTCATCTGAGCTTCGCGGAAACCATCCGCCACCTGGCTGCCGG 1037
 747 TTCCTTTCTGGACCTCGAGAGCTCCCTCCCTGATTTGACAGCAATCGGTGCCAAG 806
 1038 CGCCTTCGCTGACCTGGCGCCCTGCTGCTGCTGACCTGGATGGCAACCCGGTGAAC 1097
 807 CTTGGGGAGGACACCTCCGGGGCTGGTCAACCTCGACGACTTTCGTGAAACAA 866
 1098 ACTGGGGAGGCGAGCTGGCGGCTGGTCAACTTCGCCACATCTCTCAGCAACA 1157
 867 CCAGCTGGCGCATCCAGATGAGGCTTTTGGAGACTTCTGCTGACATTTGGAGATCT 926
 1158 CCAGCTGGCAGGCTGGCGGCGCGGCTGGATGATTTGTCGCGAGACACTGGAGGAC 1217
 927 GGACCTCTCTACAAACAACCTTCCATGGCTCCGCTGGGACTCCGTGGACCATGGTCAA 986
 1218 CGACCTCTCTACAAACAACCTCGAGCAGCTCCCTGGGAGGCGCTGGCGCCCTGGGAA 1277
 987 CCTCCACAGCTGAGCTGGACCAACCTGCTGATCACAATCGCCGAGGACACTTTTC 1046
 1278 CGTCAACAGGTTGGGCTCGACCAACCTGCTGCTTCTGTGCC---GGGCTTTTC 1334
 1047 AGACCTGCAGAAACTGGCGCGCTGGATCTCACCTTCAATCGGTGACAGAGTGCCTCC 1106
 1335 CCGCTTGCACAGCTGGCCCGCTGGACATGACCTTCAACCGCTGACCAATCCACC 1394
 1107 TGATCCCATCTTTGGCCGCTCCAGGCTTGGCTTTGACGCCACACCTTTTGGCCACC 1166
 1395 CGACCCACTCTTCTCCGCTGCTCCCTGCTCCAGGCGCGGGGCTCCCGCTCTTGC 1454
 1167 CTGTCTCTTAGTTTGGGGTAAACCCTTCACTGCAATTTGAGCTTCTCTGGCTGG 1226
 1455 CCTGGTGTGGCTTTGGCGGAAACCCCTGACTGCAACTGGAGCTGGTGGCTGG 1514
 1227 GAGGCTCGAGGGGAGATGACTGGAAAACCTGTGGCTCCAGGGGGCTCAAGGTTGG 1286
 1515 TCCTGCTGGCGGGAGGACGACCTCGAGGCTCCGCGTCCCACTGCTTGGGGGGCC 1574
 1287 CTACTTCTGGCATGCTGGAGGAGTTTGTGCGAGCGGCTTCTCATACCCAGCA 1346
 1575 CTACTTCTGGCGGTGGGAGGAGGTTTGTGCGAGCGGCTTCTGGAGCGGCTGGTCA 1634
 1347 CACACAAAGTGTGCTGTTCTGGAGGCGCAGGGCGGCACTCAAGTGCANAGCAATGG 1406
 1635 CTCACCACTCTGGCTGTGCCCCAGGCTCGGCGGCTGCTCCCTCGCTGGCGGCGAGTGG 1694

PRIOR FILING DATE: 2000-05-23
 PRIOR APPLICATION NUMBER: 60/206368
 PRIOR FILING DATE: 2000-05-23
 PRIOR APPLICATION NUMBER: 60/209832
 PRIOR FILING DATE: 2000-06-05
 PRIOR APPLICATION NUMBER: 60/218371
 PRIOR FILING DATE: 2000-07-13
 PRIOR APPLICATION NUMBER: 60/222695
 PRIOR FILING DATE: 2000-08-02
 PRIOR APPLICATION NUMBER: 60/229896
 PRIOR FILING DATE: 2000-09-01
 PRIOR APPLICATION NUMBER: 60/230621
 PRIOR FILING DATE: 2000-09-05
 PRIOR APPLICATION NUMBER: 60/232887
 PRIOR FILING DATE: 2000-09-15
 PRIOR APPLICATION NUMBER: 60/235147
 PRIOR FILING DATE: 2000-09-22
 PRIOR APPLICATION NUMBER: 60/261878
 PRIOR FILING DATE: 2001-01-12
 PRIOR APPLICATION NUMBER: 60/261910
 PRIOR FILING DATE: 2001-01-16
 PRIOR APPLICATION NUMBER: 60/261939
 PRIOR FILING DATE: 2001-01-16
 PRIOR APPLICATION NUMBER: 60/262150
 PRIOR FILING DATE: 2001-01-16
 PRIOR APPLICATION NUMBER: 60/264395
 PRIOR FILING DATE: 2001-01-25
 PRIOR APPLICATION NUMBER: 60/266421
 PRIOR FILING DATE: 2001-02-02
 PRIOR APPLICATION NUMBER: 60/267623
 PRIOR FILING DATE: 2001-02-09
 PRIOR APPLICATION NUMBER: 60/274399
 PRIOR FILING DATE: 2001-03-09
 PRIOR APPLICATION NUMBER: 60/280982
 PRIOR FILING DATE: 2001-04-03
 PRIOR APPLICATION NUMBER: 60/282129
 PRIOR FILING DATE: 2001-04-04
 PRIOR APPLICATION NUMBER: 60/282199
 PRIOR FILING DATE: 2001-04-04
 PRIOR APPLICATION NUMBER: 60/290589
 PRIOR FILING DATE: 2001-05-09
 PRIOR APPLICATION NUMBER: 09/180997
 PRIOR FILING DATE: 1998-11-19
 PRIOR APPLICATION NUMBER: 09/267213
 PRIOR FILING DATE: 1999-03-12
 PRIOR APPLICATION NUMBER: 09/380137
 PRIOR FILING DATE: 1999-08-25
 PRIOR APPLICATION NUMBER: 09/380138
 PRIOR FILING DATE: 1999-08-25
 PRIOR APPLICATION NUMBER: 09/403297
 PRIOR FILING DATE: 1999-10-16
 PRIOR APPLICATION NUMBER: 09/423741
 PRIOR FILING DATE: 1999-11-10
 PRIOR APPLICATION NUMBER: 09/709238
 PRIOR FILING DATE: 2000-11-08
 PRIOR APPLICATION NUMBER: 09/802706
 PRIOR FILING DATE: 2001-03-09
 PRIOR APPLICATION NUMBER: 09/872035
 PRIOR FILING DATE: 2001-06-01
 PRIOR APPLICATION NUMBER: 09/918585
 PRIOR FILING DATE: 2001-07-30
 PRIOR APPLICATION NUMBER: 09/924419
 PRIOR FILING DATE: 2001-08-06
 PRIOR APPLICATION NUMBER: 09/927796
 PRIOR FILING DATE: 2001-08-09
 PRIOR APPLICATION NUMBER: 09/929404
 PRIOR FILING DATE: 2001-08-13
 PRIOR APPLICATION NUMBER: 09/931836
 PRIOR FILING DATE: 2001-08-16
 PRIOR APPLICATION NUMBER: 09/941992
 PRIOR FILING DATE: 2001-08-28
 PRIOR APPLICATION NUMBER: 09/946374
 PRIOR FILING DATE: 2001-09-04

1407 GGACCCAGCCCTTTACCTGGTGGTAGCCCCCGATGACCCCGCTGGTGGAGAACTCCTC 1466
 1695 GGACCCAGACCCCGCTGGTGGTAGCCCCCGATGACCCCGCTGGTGGAGAACTCCTC 1754
 1467 AAGGACCCGCTGCTATGACAAATGGACCCCGTGGACATCTTTCATCACCACATCTCAGACAG 1526
 1755 CCGTGCAGCCGCTTCCCAATGGAGCCGTGGAGCTGCTGTCACCGAGCCGGGTGATGG 1814
 1527 TGGTGCTTCACTGCACTGTCGCAATGCTCCGGAGAGCCCGCTGGTGGAGAACTCCTC 1586
 1815 TGCGATCTTCACTGCACTGTCGCAATGCTCCGGAGAGCCCGCTGGTGGAGAACTCCTC 1874
 1587 CTCATGCTTCCAGCTGCTCCAGTTCAGCAACACAGCCAGCCAGCCAGCCAGCCAGCCAG 1644
 1875 GACTGTGGTCCCGCCACCACTTCCAGTTCAGCAACACAGCCAGCCAGCCAGCCAGCCAG 1934
 1645 TCCCGCTTCCAGCATCTGCTCCAGTTCAGCAACACAGCCAGCCAGCCAGCCAGCCAGCCAG 1704
 1935 GGACGGGATCTGATGCTTCCAGTTCAGCAACACAGCCAGCCAGCCAGCCAGCCAGCCAG 1994
 1705 GCGGAGAGCCCTCCCAAAAGCCCGGAAACCGGCTGCTTGTGTTGATGAACTGACACC 1764
 1995 CGACACTGGCCCGCTTACCAATGCTCCGGAGAGCCCGCTGGTGGAGAACTCCTC 2046
 1765 ACCTGCGCTTCAAGTGTCTGTAGCAAGTTCAGCAACACAGCCAGCCAGCCAGCCAGCCAG 1824
 2047 ACAGTGTCTTGTCCAGTGGCCGATCAGCGCCCTTCCAGTTCAGCAACACAGCCAGCCAG 2106
 1825 CTGCAAGTCTGCTGACAGTGTCTGCAAGTTCAGCAACACAGCCAGCCAGCCAGCCAGCCAG 1884
 2107 ATCCAGTCAAGAGCTGGCTGATGATCCTTCTGCTCAGAGTATCCCGGGAGAGC 2166
 1885 AAGGCCCTCGTGGTCAACAACTGTGTGCTGAGGAGTGGCTAGGACTTGTGTGTGCTGGCC 1944
 2167 CGCTGCTTCTGCTGACGGACTGGGCTCAGCGGACCTACGATCTGTGCTGCTGCC 2226
 1945 ATGTGGATGACACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2004
 2227 GTGTATGAGGAGAGCCGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2286
 2005 TCCACCAAGGCTGACTACCCAGAGTCCAGTCCAGTCCAGCAGCAGCAGCAGCAGCAGCAGCAG 2064
 2287 TCCACCAAGGCTGACTACCCAGAGTCCAGTCCAGTCCAGCAGCAGCAGCAGCAGCAGCAGCAG 2346
 2065 ATGATCCTGGTCAATCGGGGCTCATGCTGGGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2124
 2347 ATGATCCTGGTCAATCGGGGCTCATGCTGGGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2406
 2125 ATGGTGGCTAACAGGTCTGCAACACAGGAGGCCCCAG-----CAAGATGGCAGCGGCC 2178
 2407 CTAATGCGCTACAAGGTGCAACAGGAGGCCCCAG-----CAAGATGGCAGCGGCC 2466
 2179 GTGCAATGCTACTCGGAGACCAACAGGAGGCCCCAG-----CAAGATGGCAGCGGCC 2238
 2467 GTTAGCAGGCTTGTCTCCAGACCAACAGGAGGCCCCAG-----CAAGATGGCAGCGGCC 2526
 2239 GCGGGGGCCCCCGCGAGG 2257
 2527 CCGCCCCCGAGCCCGCGG 2545

APPLICANT: Stephan, Jean-Phillippe
 APPLICANT: Watanbe, Colin
 APPLICANT: Wood, William
 APPLICANT: Zhang, Zemin
 APPLICANT: Fong, Sherman
 TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
 TITLE OF INVENTION: ACIDS ENCODING THE SAME
 FILE REFERENCE: P3630RIC15
 CURRENT APPLICATION NUMBER: US/10/238,283
 PRIOR FILING DATE: 2002-09-09
 PRIOR APPLICATION NUMBER: 10/197942
 PRIOR FILING DATE: 2002-07-18
 PRIOR APPLICATION NUMBER: 60/059114
 PRIOR FILING DATE: 1997-09-17
 PRIOR APPLICATION NUMBER: 60/063046
 PRIOR FILING DATE: 1997-10-24
 PRIOR APPLICATION NUMBER: 60/065027
 PRIOR FILING DATE: 1997-11-10
 PRIOR APPLICATION NUMBER: 60/079689
 PRIOR FILING DATE: 1998-03-27
 PRIOR APPLICATION NUMBER: 60/086478
 PRIOR FILING DATE: 1998-05-22
 PRIOR APPLICATION NUMBER: 60/087607
 PRIOR FILING DATE: 1998-06-02
 PRIOR APPLICATION NUMBER: 60/089801
 PRIOR FILING DATE: 1998-06-18
 PRIOR APPLICATION NUMBER: 60/090557
 PRIOR FILING DATE: 1998-06-24
 PRIOR APPLICATION NUMBER: 60/090689
 PRIOR FILING DATE: 1998-06-25
 Remaining prior application data removed - See File Wrapper or PALM.
 NUMBER OF SEQ ID NOS: 116
 SEQ ID NO 99
 LENGTH: 2855
 TYPE: DNA
 ORGANISM: Homo Sapien
 US-10-238-283-99

Query Match 19.8%; Score 621.8; DB 9; Length 2855;
 Best Local Similarity 61.3%; Pred. No. 1.3e-162;
 Matches 1078; Conservative 0; Mismatches 662; Indels 19; Gaps 4;

QY	507	GTTCGGTGTGTCAGCGCTGCCCAAGTACTGTGTGTCAGGAAATCTGTCTGAGTCACT	566
Db	798	GTGAGCCACACCCAGCCCAATGTCCCGCGCTGCGCGTCCAGACACACAGTCTGCCCT	857
QY	567	GGGAGCCCTGCTCCCTTCAAGGGCTGCTCTTTGATACCCCTGATATTGACGGGGAC	626
Db	858	AAGCGTGTGTGTCAGCGCTGCCCAAGTACTGTGTGTCAGGAAATCTGTCTGAGTCACT	917
QY	627	AGTGGAGTTCGGCTGGGCGCAACTTTCATCATCCACATCCAGCCCGCAGGACTTTGCCAA	686
Db	918	AGCCGAGTTCGGCTGGGCGCAACTTTCATCATCCAGCCCGCAGGACTTTGCCAA	977
QY	687	CATGACGGGCTGTGGACTGACCCCTGTCAGGAAACCACTATCAGCCCAATCCAGCCCTT	746
Db	978	CATGACAGCCCTGCTGCATCTGAGCCCTGTGCGGAAACCACTATCCGCCACGTGGCTC	1037
QY	747	TTCTTTTTCGACCTCGAGCCCTCGCTCCCTGATCTTGCAGCATCCGCTCGCCAG	806
Db	1038	CGCCCTTCGCGGACTCGGCGCCCTCGCTGCTGACCTGATGACAGCCGCTGACCTC	1097
QY	807	CCTTGGGAGGACACCCCTCCGGGGCTGGTCAACCTCGAGCAGCCCTTATCGTGAACAA	866
Db	1098	ACTGGGAGGCGGAGCTGGCGGCTGGTCAACTTGGCCACTTCTCAGCAACAA	1157
QY	867	CCAGTGGCGGCACTGGCAGATGAGGCTTTTGGAGACTTCTGCTGACATGAGGATCT	926
Db	1158	CCAGTGGCGGCACTGGCGGCTGGTCAACTTGGCCACTTCTCAGCAACAA	1217
QY	927	GGACCTTCTTCAACAACACTTCATGGCTCCGCTGGGACTCCGTCGACAGCATGGTCAA	986
Db	1218	CGACCTTCTTCAACAACACTTCATGGCTCCGCTGGGACTCCGTCGACAGCATGGTCAA	1277

RESULT 11
 US-10-238-283-99
 Sequence 99, Application US/10238283
 Publication No. US20030073190A1
 GENERAL INFORMATION:
 APPLICANT: Baker, Kevin
 APPLICANT: Eaton, Dan
 APPLICANT: Filvaroff, Ellen
 APPLICANT: Goddard, Audrey
 APPLICANT: Grimaldi, J. Christopher
 APPLICANT: Gurney, Austin
 APPLICANT: Smith, Victoria

QY 987 CCTCCACGAGTGGAGCTGGACCAACAACCTGTTGTGATCAATCGCCGAGGAGACCTTTGC 1046
DB 1278 CGTAAACAGTGGGCTCGACCAACAACCTGTTGTGATCAATCGCCGAGGAGACCTTTGC 1334
QY 1047 AGACCTGAGAACTGGCCGCTGGATCTCACTCCCAATCGGCTGAGAGAGCTCCCGCC 1106
DB 1335 CGGCTGCAACAAGTGGCCGCTGGATGATGACATGACCTCCAAACCCGCTGACCAATCCAC 1394
QY 1107 TGATCCATCTTTGGCCGCTCCAGGCTTCGGCTTTTGAAGCCACACCCCTTTGCCCGCC 1166
DB 1395 CGACCCACTTCTCCCGCTGCCCCTGCTCGCCAGGCCCGGGGCTGCGCCGCTCTGC 1454
QY 1167 CTTTGTCTTTAGTTTTGGGGTAACCCACTTCACTGAAATGTAGACTTCCTGCTGCG 1226
DB 1455 CTTGGTGTCTGGCTTTGGGGAAACCCCTGCAACTGCAACTGCACTGAGTGTGCTGG 1514
QY 1227 GAGGCTGAGCGGAGGAGTACTGAAACCTTGCTCCCGAGGGGCTCAAGGGTGC 1286
DB 1515 TCGCCTGGCGGGAGGAGACCTCGAGGCTTCGCGCTCCCACTCGCTCTGGGGGCGG 1574
QY 1287 CTACTTCTGGCATGTGGTAGGAGGAGTTTGTGCGAGCCGCTCTCATCACCAGCA 1346
DB 1575 CTACTTCTGGCGGTGGGGAGGAGTTTGTGCGAGCCGCTGGTACTCACCG 1634
QY 1347 CACACAAGTGTGGTTTGGAGGCCAGGGCCACACTCAAGTGAAGGCAATGG 1406
DB 1635 CTCACCACCTCTGGCTGTGCCCCAGGTCGGCCGCTGGCTGGGGGAGTGG 1694
QY 1407 GGACCCAGCCCTTTACTACTGGGTAGCCCGGATGACCGCTGGTAGGAACTCTCT 1466
DB 1695 GGACCCAGCCCGTGTGGTGGTGTACCCAGGGCCGCTGCTAGGCAACTCAAG 1754
QY 1467 AAGGACCGCTGTATGACAAATGGCACCCCTGGACATTTTCAACACATCTCAGACAG 1526
DB 1755 CCGTGGCCGCTTTCCCAATGGACGCTGGAGTGTGTTCAACCGAGCGGGTATGG 1814
QY 1527 TGTGCTTCACTGATGCTGCAATGCTGGAGAGGAGCCAGCCGATGTTGGAGGT 1586
DB 1815 TGGCATTTTCACTGCAATGGGCAATGAGTGGAGGAGCCAGACTGCTGTGGAGCT 1874
QY 1587 CTCACCTGTCAGTGCACACCTCAGCAACAGCAGCCAGCCGCA --CTGCACCCCGCAAG 1644
DB 1875 GACTGTGGTCCCACACACTCTCAGTAGCAACAGCAGCTGTGACCCCGCGG 1934
QY 1645 TCCCGCTCTCAGACATCACTGCTCAGCAAGAACAGCCGGGAGTGGAGGAGTGG 1704
DB 1935 GGACGGGATCTGATGCTCTCAGCCCACTCCGCTGCTCTTCTGCAAGGTGG 1994
QY 1705 GCGGAGGCTCCCAAGCGCCCGGAAAGCGGCTGTGTTGCTGAAAGTGCACC 1764
DB 1995 CGACACTGGGCCCCCTA-----CCGACCGTGGCTCCAGTGTGACTGACCGGGCC 2046
QY 1765 ACCTCGGCTGTCAGTGTCTGTGAGCAAGTCAAGCAGCCCGGTTGAAGTATACAG 1824
DB 2047 ACAGTGTCTGTCCAGTGGCCGATCAGGGCTTATCCCGGCATCCCGCATGTACAG 2106
QY 1825 CTGAGTACAATGCTGTGACATGAGTACTGATTTACAGATGATCCAGCTCCAC 1884
DB 2107 ATCCAGTACAACAGTCCGCTGATGACATCTCTGCTACAGGATGATCCCGGGGAGG 2166
QY 1885 AAGGCTTCTGTGTAACAACCTGTTGAGGAGTGGCTGACTGGCTGTGTTGTTGCTGCC 1944
DB 2167 CGTCTGTTCTGTGACGGACTGGGTCAGCGGACCTTACGCTGGCTGCTGCC 2226
QY 1945 ATGTGGATGACACAGCCACCACTACAGCCACAAATCTGTTGGCTGGCCCGCAGTTC 2004
DB 2227 GTGTATGAGGACAGCCCGGCTCACGGCTCACGGCTGTTGGCTGGCCCGCTTC 2286
QY 2005 TTCACCAAGGTGACTACCCGAGTCCAGTCCATGACAGGATTTCTGGGGGAGC 2064
DB 2287 TCCACCGAACTTCGCTGGGCCATGCGGGGGCCGACAGCTCTCTGTTGGGGGAGC 2346

QY 2065 ATGATCTGCTATCGGGGGCATCATGTGGCCACCGCTGCTGTTTCATCTGTCATCTC 2124
DB 2347 ATGATCATCGGCTGGCGGGTCACTGCTAGCTCGGTCTGTTTCATCTGTTGCTG 2406
QY 2125 ATGTGGCTTACAAGTCTCAACACCAGGCGGCCCGCCAG-----CAAGATGGAGGGCC 2178
DB 2407 CTAATGCGCTTACAAGTGTCAAGCGGCGCCAGCCCGGCCCGCAGGCTCCCAAGATTC 2466
QY 2179 GTGAGCAATGTACTCGCAGACCAACGCGCCAGCCACCGCTCCCAAGCAGCGCACCA 2238
DB 2467 GTTAGCAGCGTTTGTCTCCAGACCAACGCGCCCGCTGGCGCCCGCCCGCCCG 2526
QY 2239 GCCGGGGCGCGCCGAGG 2257
DB 2527 CCCGCCCCGAGCGCCGCGG 2545

RESULT 12
US-10-238-370-99
; Sequence 99, Application US/10238370
; Publication No. US20030073191A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin
; APPLICANT: Eaton, Dan
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Goddard, Audrey
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin
; APPLICANT: Smith, Victoria
; APPLICANT: Stephen, Jean-Phillippe
; APPLICANT: Watambe, Colin
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; APPLICANT: Fong, Sherman
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3630RIC10
; CURRENT APPLICATION NUMBER: US/10/238,370
; CURRENT FILING DATE: 2002-09-09
; PRIOR APPLICATION NUMBER: 10/197942
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/059114
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/063046
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/065027
; PRIOR FILING DATE: 1997-11-10
; PRIOR APPLICATION NUMBER: 60/079689
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/086478
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090689
; PRIOR FILING DATE: 1998-06-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 116
; SEQ ID NO 99
; LENGTH: 2855
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-238-370-99

Query Match 19.8%; Score 621.8; DB 9; Length 2855;
Best Local Similarity 61.3%; Pred. No. 1.3e-162;
Matches 1078; Conservative 0; Mismatches 662; Indels 19; Gaps 4;

QY 507 GTTGGCGGTGGTCGAGCCCTGCCCAAGTACTGTGTCGCCAAGTACTGTGTCGAGTCACT 566

Db 798 GTCAGCCACACCCAGCCCATGTCCCGCGCTGCCGCTGCCAGACACAGTGCCTGCCCT 857
 Qy 567 GGGGACCTGTGCCCTCCAAGGGGCTGCTTTTGTACCCCTGATATTGACCGGGGAC 626
 Db 858 AAGCGTGTGTGCCAGGGGAGGCTCTCTGTTGTCGACACCTCGCTGGACCGCGGGC 917
 Qy 627 AGTGAGTGTGCCCTGGGGGAACTTCAATCAATCAATCAATCAATCAATCAATCAATCA 686
 Db 918 AGCCGAGTGTGCCGAGCAACTTCAATCAATCAATCAATCAATCAATCAATCAATCA 977
 Qy 687 CATGACGGGCTGTGTGACCTGACCTGTCAAGAACACCATCAGCCACATCAGCCCTT 746
 Db 978 CATGACGGCTGTGTGACCTGAGCTGTGCGGAACACCATCAGCCACATCAGCCCTT 1037
 Qy 747 TTCCTTTTGGACCTCGAGAGCTCCGCTCCCTGATCTTGACAGCAATCGCTGCCAAG 806
 Db 1038 CGCCTTTCGCGACCTGCGGGCCCTGGTGCCCTGACCTGATGGAACCGGCTGACCTC 1097
 Qy 807 CCTTGGGAGGACACCTTCGGGGGCTGTAACTGTGACGACCTTATGTTGAAACAANA 866
 Db 1098 ACTGGGCGAGGGCCAGCTGCGGGGCTGTGTAACCTTGGCCACCTCATCTCAGCAACA 1157
 Qy 867 CCAGCTGGGGCATCGCAGATGAGCTTTTGGAGACTTCTGCTGACATTTGGAGACT 926
 Db 1158 CCAGCTGGCAGGCTGGCGGGCCCTGGATGATTTGCGGAGACACTGGAGACT 1217
 Qy 927 GGACCTCTCTAACAACCTCCATGCTGCTGCTGGGACTCCGTGCGACGATGTCAA 986
 Db 1218 CGACCTCTCTAACAACCTCGAGAGCTGCTGGAGGCTTGGGAGGCTTGGGCGCTGGGCA 1277
 Qy 987 CCTCAGCAGTGTGAGCTGAGCAACAACCTGTGATCACAATCGCCGAGGCACTTTGC 1046
 Db 1278 GGTCAACACCTTGGGCTCGACCAACAACCTGTGCTGCTGCTGCTGCTGCTGCTGCT 1334
 Qy 1047 AGACCTGAGAACTGGCCGCTGATCTCACTTCACTTCACTTCACTTCACTTCACTT 1106
 Db 1335 CGCCTGCAACAGTGTGGCGGCTGGACATGACCTTCAACACCGCTGACCAACTCCACC 1394
 Qy 1107 TGATCCCATTTTGGCCGCTCCAGGCTTCCGCTTTGACAGCACACACCTTTGGCCACCC 1166
 Db 1395 CGACCCACTTCTCCCGCTGCCCTGCTCCGCGAGGCTCCCGGCTCGCCGCTCTGCTGC 1454
 Qy 1167 CTTGCTCTTTTGTGGGGTAAACCACTTCACTGCAATTTGAGCTTCTCTGCTGCG 1226
 Db 1455 CCTGCTGTGGCTTTTGGCGGAAACCCCTGCACTGCAACTGCACTGCTGCTGCTGCTG 1514
 Qy 1227 GAGGCTGAGGGAGATGACTTGAACCTTGGCTCCCGAGGGGCTCAAGGCTG 1286
 Db 1515 TCGCCTGGCGGGAGGACGACTCGAGGCTCGCGCTCCCACTGCTTGGCGGGCC 1574
 Qy 1287 CTACTTCTGGCATGTGCGTGGAGGAGTTTGTGCGAGCGGCTTCTCATCACCCAGCA 1346
 Db 1575 CTACTTCTGGCGTGGGAGGAGGTTTGTGCGAGCGGCTCCCGGCTGCTGCTGCTGCTG 1634
 Qy 1347 CACACAAAGTTGCTGGAGGCGGAGGCGGCACTCAAGTGCMAAGCCATTGG 1406
 Db 1635 CTCACACCTCTGGCTGTGCCCGGCTGCGCGGCTGCGCTGCGCTGCGCGGAGTGG 1694
 Qy 1407 GGACCCAGCCCTTATCCACTGGTAGCCCCGATGACCCGCTGGTAGGAACTCTCTC 1466
 Db 1695 GGACCCAGAGCCCGTGTGGGCTGTCACCCAGGCGGCTGTAGGCAACTCAAG 1754
 Qy 1467 AAGGACCGCTGTATGACAAATGGACCTGGACATCTTCAACCCACATCTCAGACAG 1526
 Db 1755 CCGTCCCGCGCTTCCCAATGGAGCTGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1814
 Qy 1527 TGGTGCCTTCACTGCAATGCTCCAGGAGGCGGCAACCGGCAATGGTAGGT 1586
 Db 1815 TGGCATCTTCACTGCAATGCTCCAGGAGGCGGCAACCGGCAATGGTAGGT 1874
 Qy 1587 CTCATGCTCAGCTGCACTCAGCAACAGCAGCCAGCCGCA--CTGACCCCCCAAG 1644
 Db 1875 GACTGTGGTCCCGCCACCACTCTCAGCTAGTCCCAACAGCACCAGCTGTGACCCCGCGG 1934

Qy 1645 TCCCGCTTCTCAGACATCACTGGTTCAGCAAGACACCGCGGGAGGTGGAGGCACTGGG 1704
 Db 1935 GGACGGGATCTGATGTCTTCAACCCACCTCCGCTGCTGCTTCTGCCAAGTGGC 1994
 Qy 1705 GCGGAGAGCTTCCAAAAGCCCCCGGAAAGGGCTGTGCTTGTGCTGAAAGTACCACC 1764
 Db 1995 CGACACTGGGCCCCCTA-----CCGACCGTGGCGTCCAGGTGACTGAGCACGGGGCC 2046
 Qy 1765 ACCTCGGCTTGGTCAAGTGGTCTGTACAGAACTGACAGCCAGCCCGGTTGAAGATGTACCAG 1824
 Db 2047 ACAGTGTCTTGTTCAGTGGCCGATCAGGGCTATCCCGGGCATCCGCAATGTACCAG 2106
 Qy 1825 CTGCAATCAACTGCTTGCAGATGAGTACTGATTTACAGATGATCCAGCCTTCAAAC 1884
 Db 2107 ATCCAGTACAACAGCTCGCTGATGACATCTCTGCTACAGGATGATCCCGGGAGAGC 2166
 Qy 1885 AAGGCTTGTGGTCAACAACCTGCTGAGGACTGAGTGTGCTGCTGCTGCTGCTGCTG 1944
 Db 2167 CGTCTGTTCTGTGACGGACCTGGCGTCAAGCCGAGCTTGTGCTGCTGCTGCTGCTG 2226
 Qy 1945 ATGTTGGATGACACAGCCAGCAGACTCAGGGCCACAACATCGTGGGCTGCGCCAGTTT 2004
 Db 2227 GTGATGAGACAGCCAGCGGCTCAGGGCTCAGGGCCAGCGGCTGTGGCTGCGCCGCTT 2286
 Qy 2005 TTCACCAAGGCTGACTACCCCGCAGTCCAGTCCATGACAGCCAGATTTCTGGGCGGACC 2064
 Db 2287 TCCACCGAACCCTCGCTGCGCCATGCGGGGCGCCAGCGCTCCCTTCTGCGGCGCAGC 2346
 Qy 2065 ATGATCTGCTCATCGGGGCAATCATGTTGGCCAGCTGCTGGTCTTCAATCGTCACTTC 2124
 Db 2347 ATGATCATCGGCTGGGGGCTCATCGTAGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2406
 Qy 2125 ATGTTGGCTCAAGTCTGCAACCCAGGCGCCAG-----CAAGATGCGAGCGGCG 2178
 Db 2407 CTATGCGCTACAGGTGACGGCGGCGAGCCCGGCGGCAAGATTTCCGGGCT 2466
 Qy 2179 GTGACAAATGTACTCGCAGAACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2238
 Db 2467 GTTAGCAGGTTTGTCTCCAGACCAACCGGCGGCGGCGGCGGCGGCGGCGGCGG 2526
 Qy 2239 GCGGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2257
 Db 2527 CCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2545

RESULT 13
 US-10-245-055-99
 ; Sequence 99, Application US/10245055
 ; Publication No. US20030073192A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Baker, Kevin
 ; APPLICANT: Eaton, Dan
 ; APPLICANT: Filvaroff, Ellen
 ; APPLICANT: Goddard, Audrey
 ; APPLICANT: Grimaldi, J. Christopher
 ; APPLICANT: Gurney, Austin
 ; APPLICANT: Smith, Victoria
 ; APPLICANT: Stephan, Jean-Phillippe
 ; APPLICANT: Watane, Colin
 ; APPLICANT: Wood, William
 ; APPLICANT: Zhang, Zemin
 ; APPLICANT: Fong, Sherman
 ; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
 ; TITLE OF INVENTION: ACIDS ENCODING THE SAME
 ; FILE REFERENCE: P36301C88
 ; CURRENT APPLICATION NUMBER: US/10/245,055
 ; CURRENT FILING DATE: 2002-09-16
 ; PRIOR APPLICATION NUMBER: 10/197942
 ; PRIOR FILING DATE: 2002-07-18
 ; PRIOR APPLICATION NUMBER: 60/059114
 ; PRIOR FILING DATE: 1997-09-17
 ; PRIOR APPLICATION NUMBER: 60/063046

1455 CCTGGTCTGGCCCTTTGGCGGGAACCCCTTGCACTGCAACTGGAGCTGTGTGGCTGCG 1514
 1227 GAGCTCGAGCGGACGATGACCTGGAACCTGTGGCTCCCGAGGGGGCCCTCAAGGGTGC 1286
 1515 TCGCTTGGCGGGGAGGACGACCTCGAGGCTGGCGTCCACCTGCTCTGGGCGCGCC 1574
 1287 CTACTTCTGGCATGTGGCTGAGGAGGATTTGTGGAGCCGCTCTCATCACCACCA 1346
 1575 CTACTTCTGGCGGTGGCGAGGAGGATTTGTGGAGCCGCTCTCATCACCACCA 1634
 1347 CACACACAAGTTCTGGTTTGGAGGCGGAGGCGGACACTCAAGTGCAGAAAGCAATTGG 1406
 1635 CTCACCACTTGGCTGTGGCGGAGGCTGGCGGCTGGCCCTGGCTCCCGGAGTGG 1694
 1407 GGACCCCGCCCTTATCCACTGGTAGCCCGGATGACCGCCCTGGTAGGAATCTCTC 1466
 1695 GGACCCAGAGCCCGTGGCTGGTGGTGTACCCCGGCGGCTGTAGGCAACTCAAG 1754
 1467 AAGGACCGCTGTATGACAAATGGCACCCCTGGACATTTCACTACCAACATCTCAGAC 1526
 1755 CCGTCCCGCGCCCTTCCCAATGGGACGCTGGAGCTGTGTACCCAGCGGCTGTGG 1814
 1527 TGGTCCCTTACCTGCAATGCTGCAATGCTGCGGAGGCGGACCGGCAATGTTGGAGGT 1586
 1815 TGGCATCTTACCTGCAATGCGGCAATGCACTGGCGAGGCGGACAGCTGTGTGGAGCT 1874
 1587 TCCCATCGTCCAGCTGCCACACTCAGCAACAGCACAGCCAGCCGCA--CTGACCCCCCAAG 1644
 1875 GACTGTGGTCCCGCCACCTCTCAGTAGCAACAGCACAGCAGCTGTGACCCCGCCG 1934
 1645 TCCCGCTCTCAGACATCAGTGGTCCAGCAAGACAGCCGCGGAGGTGGAGCGAGTGG 1704
 1935 GGACGGGATCTGATGTCTTACCCCACTCCGCTGCTGTCTTGTGCAAGGTTGGC 1994
 1705 GGCGAGAGCTCCCAAAAGCCCGGAAAGCGGCTGTGTTGTCTGAAAGTGACACC 1764
 1995 GCACACTGGGCCCCCTA-----CCGACCGTGGCGTCAAGTGACTGACGACGGG 2046
 1765 ACCTCGGCTGTGCAAGTGTGTGACAAAGTGTGACAGCCCGGGTGAAGATGACAG 1824
 2047 ACAGTGTCTTGTCCAGTGGCGGATCAGCGGCTATCCCGGGCATCCGATGTACAG 2106
 1825 CTGAGTACAATCTCTGACGATGAGTACTGATTTACAGATGATCCAGCTCCCAAC 1884
 2107 ATCCAGTACAAGCTCGGCTGATGACATCTCTGTCTACAGATGATCCCGGGAGAGC 2166
 1885 AAGGCTTCTGGTCAACAACTGGTGTGAGGACTGGCTACGACTTGTGTGTGGCTC 1944
 2167 CGCTGTTCTGCTGACGACCTGGCTCAGCGGACTACGATCTGTGCTGTGCTGCC 2226
 1945 ATGTGGATGACACAGCACGACACTCAGGGCCCAACACTGTTGGCTGGCCAGTTC 2004
 2227 GTGTATGAGGACAGCCCGGGCTCAGGGCCAGCGGCTGTGGGCTGGCCCGCTTC 2286
 2005 TTCACCAAGGCTGACTACCCGAGTCCAGTCCATGACAGCCAGATTCGGGCGGCA 2064
 2287 TCCACCGAACCTGCGGTGGCCATGGGGGCGCCGACGCTCCCTTCTTGGGCGGACG 2346
 2065 ATGATCTGGTCAATGGGGGCAATCATGTGGCCAGCGCTGTGGTCTTCAATGTCCTC 2124
 2347 ATGATCATCGCTGGGCGCTCATCTGAGCTCGTACTGGTCTTCACTTCTGTGCTG 2406
 2125 ATGGTGGCTTCAAGGTGTGCAACCAAGAGGCGCCAG-----CAAGATGGCAGCGCC 2178
 2407 CTAATCGCTTACAAGGTGACAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGCT 2466
 2179 GTGAGCAATGTACTCGCAGAACCAAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 2238
 2467 GTTAGCAGGTTTGTCTCCAGCAGCAAGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2526
 2239 GCCGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2257
 2527 CCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2545

Db 1455
 Qy 1227
 Db 1515
 Qy 1287
 Db 1575
 Qy 1347
 Db 1635
 Qy 1407
 Db 1695
 Qy 1467
 Db 1755
 Qy 1527
 Db 1815
 Qy 1587
 Db 1875
 Qy 1645
 Db 1935
 Qy 1705
 Db 1995
 Qy 1765
 Db 2047
 Qy 1825
 Db 2107
 Qy 1885
 Db 2167
 Qy 1945
 Db 2227
 Qy 2005
 Db 2287
 Qy 2065
 Db 2347
 Qy 2125
 Db 2407
 Qy 2179
 Db 2467
 Qy 2239
 Db 2527

Query Match 19.8%; Score 621.8; DB 9; Length 2855;
 Rest Local Similarity 61.3%; Pred. No. 1.3e-162;
 Matches 1078; Conservative 0; Mismatches 662; Indels 19; Gaps 4;

Qy 507 GTTTCGGTGTGCGAGCCTGCCCCAAAGTACTGTGTGCCAGAACTGTCTGAGTCACT 566
 Db 798 GTGAGCACACCCAGCCCATGTCCCGCGCTGCCGCTGCAGACACAGTCTGCTGCCCT 857
 Qy 567 GGGACCCCTGCCCCCAAGGGTGTCTTTGTACCCCTGTATGACCGGGGAG 626
 Db 858 AAGCGTGTGTCGCCAGGGCAGGCCTCCCTGTTGTTGTCACCCCTGCTGAGCCCGGGC 917
 Qy 627 AGTGGAGCTCGCTGGCGGCAACTTCATCATCCACATCAGCCCGCAGGACTTTGCCAA 686
 Db 918 AGCCGAGCTCGGCTGGCAGACAACTTCATCGCCTCGGTGGCGCGGCGGAGCCTGGCCAA 977
 Qy 687 CATGACGGGCTGTGAGACCTGACCTGTCAGGAAACCATCAGCCACATCCAGCCCTT 746
 Db 978 CATGACAGGCTGTGCACTGAGCCTGTCGGGAAACCATCCGGCCAGTGGCTGCCGG 1037
 Qy 747 TTCCTTTCGACTCGAGACCTCGCTCCCTGCACTTTGACAGCAATCGGCTGCCAA 806
 Db 1038 CGCCTTCGCGACTCGGGCCCTGCGTGCCTGACCTGATGGCAACCGGCTGACCTC 1097
 Qy 807 CFTTGGGAGGACACCTTCGCGGCGCTGGTCAACCTGACACCTGATGGTGAACAA 866
 Db 1098 ACTGGCGAGGCGCAGCTGGCGGCGCTGGTCAACTTGGCCACTCATCCAGCAACAA 1157
 Qy 867 CCAGCTGGCGGCACTCCAGATGAGGCTTTGAGGACTTCTGCTGACATTTGAGGATCT 926
 Db 1158 CCAGCTGGCAGCTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1217
 Qy 927 GGACTCTCTTACAACAACTTCAATGCGCTGCGGCGGACTCGCTGGAGCGCATGTTCAA 986
 Db 1218 CGACTTCTTCAAAACAACTGAGCAGCTGCTGGAGGCGGCGGCGGCGGCGGCGGCGG 1277
 Qy 987 CCTCCACAGCTGAGCTGAGCAACAACTGCTGGATCAATCGCCGAGGGCAGCTTTGC 1046
 Db 1278 CGTCAACACGTTGGGCGCTCGAACCAACCTGCTGGCTTCTGTGGCC--GGCGCTTTTC 1334
 Qy 1047 AGACTCAGAAATGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1106
 Db 1335 CCGGCTGCAAGCTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1394
 Qy 1107 TGATCCCATTTTCCCGCTCCCAAGGCTTGGCTTTGACAGCCACACCTTTTGGCCCA 1166
 Db 1395 CGACCACTTCTCCCGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1454
 Qy 1167 CTTGTCTTTAGTTTGGGCTTACCACCTTCACTGCAATTTGTGAGCTTCTCTGGCTGCG 1226

SEI ID NO 99
 Remaining Prior Application data removed - See File Wrapper or PALM.
 NUMBER OF SEQ ID NOS: 116
 LENGTH: 2855
 TYPE: DNA
 ORGANISM: Homo Sapien
 US-10-245-055-99

RESULT 14

US-10-245-147-99
Sequence 99, Application US/10245147
Publication No. US20030073193A1

GENERAL INFORMATION:
APPLICANT: Baker, Kevin
APPLICANT: Eaton, Dan
APPLICANT: Filvaroff, Ellen
APPLICANT: Goddard, Audrey
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Austin
APPLICANT: Smith, Victoria
APPLICANT: Stephan, Jean-Phillippe
APPLICANT: Watanabe, Colin
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
APPLICANT: Fong, Sherman

TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

FILE REFERENCE: P3630R1C72

CURRENT APPLICATION NUMBER: US/10/245,147

PRIOR FILING DATE: 2002-09-16

PRIOR APPLICATION NUMBER: 10/197942

PRIOR FILING DATE: 2002-07-18

PRIOR APPLICATION NUMBER: 60/059114

PRIOR FILING DATE: 1997-09-17

PRIOR APPLICATION NUMBER: 60/063046

PRIOR FILING DATE: 1997-10-24

PRIOR APPLICATION NUMBER: 60/065027

PRIOR FILING DATE: 1997-11-10

PRIOR APPLICATION NUMBER: 60/079689

PRIOR FILING DATE: 1998-03-27

PRIOR APPLICATION NUMBER: 60/086478

PRIOR FILING DATE: 1998-05-22

PRIOR APPLICATION NUMBER: 60/087607

PRIOR FILING DATE: 1998-06-02

PRIOR APPLICATION NUMBER: 60/089801

PRIOR FILING DATE: 1998-06-18

PRIOR APPLICATION NUMBER: 60/090557

PRIOR FILING DATE: 1998-06-24

PRIOR APPLICATION NUMBER: 60/090689

PRIOR FILING DATE: 1998-06-25

Remaining prior Application data removed - See File Wrapper or PALM.

NUMBER OF SEQ ID NOS: 116

SEQ ID NO 99

LENGTH: 2855

TYPE: DNA

ORGANISM: Homo Sapien

US-10-245-147-99

Query Match 19.8%; Score 621.8; DB 9; Length 2855;

Best Local Similarity 61.3%; Pred No. 1.3e-162;

Matches 1078; Conservative 0; Mismatches 662; Indels 19; Gaps 4;

507 GTTTCGCGGTGCGAGCCGCTGCCCAAGTACTGTGTCTGCCAGAATCTGTCTGAGTCACT 566

798 GTGAGCAGCCAGCCAGCCATGTCCTCCCGCGCTGCCAGACACAGTCTGCTCCCT 857

567 GGGGACCTGTGCCCTCCAGGGGCTGTCTTTGTACCCCTGATATTTGACCGGGGGAC 626

858 AAGCGTGTGTGCCCGAGGGGCTCCTGTTCGTGCCACCCCTCGCTGCCGCGGGGCG 917

627 AGTGAGCTGCGCCTGGGGGCAACTTTCATATCCACATCAGCCCGCAGGACTTTGCCAA 686

918 ACCCGAGCTGCGGCTGGCAGAACATTCATCGCTTCCTGCGCCCGCGGCACTGGCCAA 977

687 CATGACGGGCTGGTGGACCTGACCCCTGCCAGAACCATCAGCCACATCAGCCCTT 746

978 CATGACAGGCTGTGCATCTGAGCTGTCCGGAAACCATCCGCCACGCTGGCTGCCGG 1037

747 TTCCTTTCTGGACCTCGAGAGCCCTCGCTCCCTGCACTTTGACAGCAAFCGGCTGCCAAG 806

Table with columns for Db, Qy, and sequence alignment. Rows 1038-1894 show sequence matches between Db and Qy with vertical bars indicating alignment positions.

Remaining Prior Application data removed - See File Wrapper or PALM.

```

; NUMBER OF SEQ ID NOS: 116
; SEQ ID NO 99
; LENGTH: 2855
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-245-730-99

Query Match          19.8%; Score 621.8; DB 9; Length 2855;
Best Local Similarity 61.3%; Pred. No. 1.3e-162;
Matches 1078; Conservative 0; Mismatches 662; Indels 19; Gaps 4;

Oy 507 GTTTGCCGTGGTTCGACGCTGCCCCAAGTACTGTGTCTGCCAAGTCTGCTCTGAGTCACT 566
Db 798 GTGACCCACACCCAGCCATGTCCTCCCGCGCTCCCGTCCAGACACAGTTCGTGCCCCCT 857
Oy 567 GGGGACCTGTGCCCTTCCAAAGGGGTGCTTTGTATACCCCTGATATTCACCGGGGGAC 626
Db 858 AAGCGTCTGTGCCCCAGGGGCGGCTTCCTGTGTCACCTCTCCCTGTCAGCCGCGGCG 917
Oy 627 AGTGGAGCTGGCTGGCGGCAACTTCATCCATCATCCAGCCGCGAGACTTTTGCCAA 686
Db 918 AGCCAGCTGGGCTGGCAGACAATTCATCGCTCCGTGGCGCCGCGCAGCTTGGCCAA 977
Oy 687 CATGACGGGCTGTGTGACCTGACCTGTCCAGGAACACCATCAGCCACATCCAGCCCTT 746
Db 978 CATGACAGGCTGTGTGACTGTGAGCTGTCCGGGAAACCATTCGCCCTCCAGCTGCTGCGG 1037
Oy 747 TTCCTTTCTGACCTCGAGAGCTCCGCTCCCTCATCTTACAGCAATCGGTGCGCAAG 806
Db 1038 CGCTTTCGCCACTGGGGCCCTCGTGGCTGACCTGATGGCAACCGGCTGACTC 1097
Oy 807 CCTGGGAGGACACCCCTCGGGGCTGGTCAACTGAGCACCTTATTCGTGACACAA 866
Db 1098 ACTGGGAGGGCAGCTCGCGGCTGGTCAACTTGGCCACCTCATCTTCAGCAAA 1157
Oy 867 CGACTGGCGGCTCCAGATGAGCTTTTGTAGACTTCTGCTGACATTTGGAGGATCT 926
Db 1158 CGAGCTGGCAGGCTGGGGGCGGCTCCGTGATGTGCCGAGACACTGGAGGACT 1217
Oy 927 GGACTCTCTTACAAACCTCCATGGCTCCGCTGGGACTCCGTCGACGATGTTGTA 986
Db 1218 CGACCTCTCTACAAACCTCGAGCAGCTGCCCTGGGAGGCTTGGGCGGCTGGGCAA 1277
Oy 987 CCTCACAGCTGAGCTGGACCAACCTGCTGATCACAATCGCCGAGGAGGACCTTTGC 1046
Db 1278 CGTCAACAGTTGGGCTCGACCAACCTGCTGGCTTCTGTGCCCC---GGCGCTTTTC 1334
Oy 1047 AGACTGCAAACTGGCCGCTGGATCTCACTCCAAATCGGCTGCAAGACTCCCC 1106
Db 1335 CGGCTGCAAAAGTGGCCGCTGGACATGACTTCAACCGCTGACCAATCCACC 1394
Oy 1107 TGATCCATCTTTGGCCGCTCCAGGCTTCGCGTTTGACAGCCACCTTTGCCCCACC 1166
Db 1395 CGACCCACTTCTCCGCTGCTCCGCTGCTCCAGGCGCCCGGGCTCCGCCCTTGC 1454
Oy 1167 CTTGCTCTTGTGTTGGGGTAAACCACTTCACTGCAATTTGTAGCTTCTCTGGCTGG 1226
Db 1455 CTTGCTCTTGTGTTGGGGTAAACCACTTCACTGCAATTTGTAGCTTCTCTGGCTGG 1514
Oy 1227 GAGGCTCGAGGGGAGGAGTGCATGGAACCTGTGCTCCCGAGGGGGCTCAAGGGTGC 1286
Db 1515 TCGCTGGGCGGAGGAGGAGCTCGAGGCTCGCGGCTCCCACTGCTCGGGCGGCG 1574
Oy 1287 CTACTTCTGGCATGTGCTGAGGAGGAGTTGTCGAGCGGCTCTCATCAGCCAGCA 1346
Db 1575 CTACTTCTGGCGGTTGGGAGGAGTTGTCGAGCGGCTCGCGGCTCGGGCGGCG 1634
Oy 1347 CACACAGTTCCTGTTCTGAGGCGGAGGCGGCGCACACTCAAGTGCAAAGCCATTTG 1406
Db 1635 CTACCACCTTGGCTGTGCCCCAGGTGGGCGGCTGGCTGCGCTGCGGGCAGTGGG 1694
Oy 1407 GGACCCCGCCCTTATCCACTGGGTAGCCCCCGGATGACCGCTGGTAGGAACTCCTC 1466

```

```

Db 2107 ATCCAGTACAACAGCTCGGCTGATGACATCTCTGTCTACAGGATGATCCCGCGGAGAGC 2166
Oy 1895 AAGGCTTTCGTTGTCACAACTGTTGTCAGGACTGGCTACGACTTGTGTGTCTGGCC 1944
Db 2167 CCCTCGTTCCTGCTGACGACCTGGCTGAGGCGGACTAGGATCTGTGGTGTCTGCC 2226
Oy 1945 ATGTGGATGACACAGCCAGACACTCAGCCACCAACATCGTGGGCTGGCCAGTTC 2004
Db 2227 GGTATGAGACAGGCCACGGGCTCAGGGCAGCCGCTGTGGGCTGGCCCGCTTC 2286
Oy 2005 TTCACAGGCTACTACCCGAGTCCAGTCCATGACAGCCAGACTTCGGGCGGACC 2064
Db 2287 TCCACCGAACCTGCGCTGGGCGCATGCGGGGCGCCGACGCTCTCTTGGCGGACG 2346
Oy 2065 ATGATCTCGTTCATCGGGGCTCATCTGTGGCAGCTGCTGGTCTTTCATCGTATCTC 2124
Db 2347 ATGATCATCGCTGGGCGGCTCATCTGTAGCTCGGTACTGTTCTCATCTCTGCTG 2406
Oy 2125 ATGGTCCGTACTAAGTCTGACACACAGAGGCCCCAG-----CAAGTGGCAGGGCC 2178
Db 2407 CTAATGCGCTACAAGGTGACAGCGGCGCCAGCCCCCGGCAAGGCTCCCGGCT 2466
Oy 2179 GTGAGCAATGTACTCGAGACCAACGGCGCCAGCCAGCCGCTCCAGAGCGCAGCA 2238
Db 2467 GTTAGCAGCGTTTGTCTCCAGCAACGGCGCCCTGGGCGCCCGCCAGCCCGCG 2526
Oy 2239 GCCGGGCGCCCGCGGAGG 2257
Db 2527 CCCGCCCGGAGCCCGGG 2545

RESULT 15
US-10-245-730-99
; Sequence 99, Application US/10245730
; Publication No. US20030073194A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin
; APPLICANT: Eaton, Dan
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Goddard, Audrey
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Phillippe
; APPLICANT: Watanabe, Colin
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; APPLICANT: Fong, Sherman
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3630R1C8S
; CURRENT APPLICATION NUMBER: US/10/245,730
; CURRENT FILING DATE: 2002-09-16
; PRIOR APPLICATION NUMBER: 10/197942
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/059114
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/063046
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/065027
; PRIOR FILING DATE: 1997-11-10
; PRIOR APPLICATION NUMBER: 60/079689
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/086478
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090589
; PRIOR FILING DATE: 1998-06-25

```



```

Db 1695 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1467 AAGGACCGGTGTATGCAATGGACACCTGGACATCTTATCACCACATCTCAGGACAG 1526
Db 1755 CCGTGCCTGGCTTCCCAATGGAGCGTGGAGTGTGGTACCGAGCCGGGTGATGG 1814
QY 1527 TGGTCCCTTACCTGATGTCCTGCAATGTCCTGGAGAGCCACGGCCATGGTGGAGGT 1586
Db 1815 TGGCATCTTACCTGATGTCCTGCAATGTCCTGGAGAGCCACAGCTGCTGTGGAGCT 1874
QY 1587 CTCCATCTGTCAGCTCAGCACTCAGCAACAGCACACCGCCCA--CTGCACCCCAAG 1644
Db 1875 GACTGTGGGTCCCAACACCTCTCAGCTAGCTAGCAACAGCACAGCTGTGACCCCCCGG 1934
QY 1645 TCCCGCCTCTCAGACATCACTGGCTCCAGCAAGACAGCCGGGAGGTGGAGCCAGTGGG 1704
Db 1935 GGACGGGATCTGTATGCTCTCACCCCACTCCCGCTGCTGCTTCTGCCAAGGTGGC 1994
QY 1705 GCGGAGAGCTCCCAAGCCCGGAAACGGGTGTGCTGTGTGTAAGTACACC 1764
Db 1995 CGACACTGGGCCCTA-----CCGACCTGGGCTCAGGTGACTGAGCACGGGGCC 2046
QY 1765 ACCTGGCCCTGGTCAAGTGTCTGTCAGCAAGTACAGCACCCCGGTGAGATGTACCAG 1824
Db 2047 ACAGTGTCTTGTCCAGTGGCCGATCAGCGGCTATCCCGGATCCGCAATGATACCAG 2106
QY 1825 CTGCAGTAACTGTCTGACGATGAGTGTGATTTACAGGATGATCCAGCCCTCCAC 1884
Db 2107 ATCCAGTAAACAGTCCGCTGATGACATCTCTGCTACAGGATGATCCCGGGAGAGC 2166
QY 1885 AAGGCTTGTGGTCAACAACTGTTGTGAGGACTGGCTACGACTGTTGTGTGTGGCC 1944
Db 2167 CGCTGCTTCTGCTGACGACCTGGCTGAGCCGGACCTACGATCTGTGCTGTGCTG 2226
QY 1945 ATGTGGATGACACAGCCAGACTACGCGCCCAACATCGTGGGCTGCGCCAGTTC 2004
Db 2227 GTGTATGAGGACAGCCACGGGCTCACGGCCACGGCCCTGTGGGCTGCGCCCGCTTC 2286
QY 2005 TCCACCAAGCTGACTACCCCGAGTCCAGTCCATGCAGCCAGATCTGGCGGCACC 2064
Db 2287 TCCACCGAACCCTGCGCTGGGCTGAGGCGCCGACGCTCCCTTCTGGCGGCAGG 2346
QY 2065 ATGATCCTGGTATCGGGGCAATCGTGGCCAGCTGCTGTTTATCGTCAATCCTC 2124
Db 2347 ATGATCATCGCTGGGCGGCTCATCGTAGCTCGGTACTGGTCTTTCATCTTGTGCTG 2406
QY 2125 ATGGTCCGCTACAGGTCTGCAACCAAGAGCCCGCCAG-----CAAGATGGCAGGGCC 2178
Db 2407 CTAATGCGCTACAAGGTGACCGCGGCCAGCCCGGCAAGGCCAAGATTCGCGGCT 2466
QY 2179 GTGACCAATGTACTCGAGACCAACAGCGCCAGCCAGCCGCTCRAAGCAGGCCCA 2238
Db 2467 GTTAGCAGGCTTGTCTCCAGACCAACAGCGGCCCTGGGCCCCACGCCACGGCCCGC 2526
QY 2239 GCCGGGCCCGCCGAGG 2257
Db 2527 CCGCGCCGAGCCCGG 2545

```

Search completed: June 8, 2003, 21:27:39
Job time : 453 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 8, 2003, 15:32:52 ; Search time 3910 Seconds
(without alignments)
13022.662 Million cell updates/sec

Title: US-09-831-846-1
Perfect score: 3144
Sequence: 1 gcctggctccctctcgtgga.....9cggcttggattgcttatg 3144

Scoring table: IDENTITY_NUC
Gapop 10_0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- 1: em_estba:*
- 2: em_esthum:*
- 3: em_estin:*
- 4: em_estmu:*
- 5: em_estov:*
- 6: em_estpl:*
- 7: em_estro:*
- 8: em_htc:*
- 9: gb_est1:*
- 10: gb_est2:*
- 11: gb_htc:*
- 12: gb_est3:*
- 13: gb_est4:*
- 14: gb_est5:*
- 15: em_estfun:*
- 16: em_estom:*
- 17: gb_gss:*
- 18: em_gss_hum:*
- 19: em_gss_inv:*
- 20: em_gss_pln:*
- 21: em_gss_vrt:*
- 22: em_gss_fun:*
- 23: em_gss_mam:*
- 24: em_gss_mus:*
- 25: em_gss_other:*
- 26: em_gss_pro:*
- 27: em_gss_tod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	2170.6	69.0	3204	11 AK017594	Mus muscu
2	609.4	19.4	611	14 BM787163	K-EST0066
3	606.8	19.3	621	12 BF312768	601896906
4	582	18.5	582	14 BQ126354	il16ell.y
5	540.2	17.2	578	14 BQ126066	il16ell.x
6	525.6	16.7	652	14 BM948451	UI-M-EGOP

7	514.2	16.4	655	10	BB642310
8	492.6	15.7	1853	11	BC007718
9	441.8	14.1	457	10	AW163387
10	422.8	13.4	435	10	AW161566
11	377.2	12.0	685	10	BB278209
12	346.6	11.0	357	13	BM193100
c	327	10.4	635	12	EG380772
13	313	10.0	397	10	AW161909
14	310.6	9.6	708	13	BM046297
15	300.4	9.6	714	14	BM950379
16	293.8	9.3	674	14	BM722122
17	293.8	9.3	674	14	BM722122
18	285.4	9.1	338	10	AW415047
19	279	8.9	563	10	AW274912
20	276.6	8.8	712	14	BM944080
21	273.4	8.7	822	14	BQ571534
22	271.4	8.6	703	14	BM784343
23	267.4	8.5	472	12	BF563759
24	266.2	8.5	512	9	AI383513
25	266.2	8.5	513	9	AI418838
26	264.8	8.4	526	10	AW003909
27	264.4	8.4	720	14	BM948138
28	263.2	8.4	635	14	BM742719
29	263	8.4	513	9	AI760811
30	260.6	8.3	634	9	AI830569
31	259.8	8.3	618	14	BM744495
32	259.8	8.3	620	14	BM742162
33	259.8	8.3	620	14	BM742315
34	259	8.2	1091	13	BM552976
35	258.2	8.2	609	14	BM743747
36	257.6	8.2	609	14	BM742740
37	257.4	8.2	751	12	BE879258
38	252.4	8.0	602	14	BM741814
39	252	8.0	618	14	BM744466
40	249.6	7.9	541	12	BF411974
41	249.6	7.9	589	14	BM743945
42	245.4	7.8	600	14	BM949682
43	243.8	7.8	581	14	BM783370
44	243.6	7.7	526	9	AA527134
45	243	7.7	637	14	BM784328

ALIGNMENTS

AK017594 3204 bp mRNA linear HTC 19-JAN-2002
Mus musculus 8 days embryo whole body cDNA, RIKEN full-length
enriched library, clone:573042005;homolog to KIAA1246 PROTEIN
(FRAGMENT), full insert sequence.

AK017594 GI:12856913
HTC; CAP trapper.

Mus musculus (strain:C57BL/6J) 8 days embryo cDNA to mRNA,
clone lib:RIKEN full-length enriched mouse cDNA library
clone:5730420005.

ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.

REFERENCE
1
AUTHORS Carninci, P. and Hayashizaki, Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636

REFERENCE
2
AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE 20499374
PUBMED 11042159

QY 491 TAGCGTTGGCAATGGCGTTTGGCGTGGTTCGACGCGCTGCCCCAAGTACTGTGTCTGCCAGA 550
 Db 491 TGGCTTTTGGCAATGGCGTTTGGCGTGGTTCGACGCGCTGCCCCAAGTACTGTGTCTGCCAGA 550
 QY 551 ATCTGTCTGAGTCACTGGGACCCCTGTGCGCCCTCCAAAGGGCTGTCTTTGTACCCCTTG 610
 Db 551 ATCTGTCTGAGTCACTGGGACCCCTGTGCGCCCTCCAAAGGGCTGTCTTTGTGCCCCCTG 610
 QY 611 ATATTGACCGGGGACAGTGGAGCTGCGCTGGGGGCAACTTTCATCATCCACATCAGCC 670
 Db 611 ACATTGACCGGGGACAGTGGAGCTGCGCTGGGGGCAACTTTCATCATCCACATCAGCC 670
 QY 671 GCCAGACTTTGCCAAATGACGCGGCTGGTGGACTGACCCCTGTCAGGAACACATCA 730
 Db 671 GCCAAGACTTTGCCAAATGACGCGGCTGGTGGACTGACCCCTGTCAGGAACACATCA 730
 QY 731 GCCACATCCAGCCCTTTCCTTCTCGGACCTCGAGCCCTCGCTCCCTGCACTTTGACA 790
 Db 731 GCCATATCCAGCCCTTTCCTTCTCGGACCTCGAGCCCTCGCTCCCTGCACTTTGACA 790
 QY 791 GCAATCGGTGCAAGCTTTGGGAGGACACCCCTCCGGGGCTGGTCAACTGCGACCC 850
 Db 791 GCAACCGGTACTTACCTTTGGGAGGACACACTCCGGGGTCTGGTCAACTGCGACCC 850
 QY 851 TTATCGTGAACAACAAACAGCTGGGGGCAATCGCAGATGAGGCTTTTGGAGACTTCTTGC 910
 Db 851 TTATTTGTGAACAATAACCAAGCTGGGTGGCACTCGCTGATGATGCTTTTGAAGACTTCTTGC 910
 QY 911 TGACATTTGGAGGATCTGGACCTTCTCTACAAACCTTCCAGCCCTGGCTGGGACTCCG 970
 Db 911 TGACTTTTGAAGACTTTGGACCTTCTTACAAACCTTCCAGCCCTGGGACTTCCCTGGGATCCG 970
 QY 971 TCGACGATGTCACCTTCCACAGCTGAGCTGGACCAACCTGCTGGATCAATCG 1030
 Db 971 TACGGGCAATGTCACCTTCCATCAGCTGAGTCTGGACCACAACTGTGTGGACCAATG 1030
 QY 1031 CCGAGGCACTTTTCAGACCTGCGAATACTGGCCCGCTGGATCTCACTTCCATTCGGC 1090
 Db 1031 CTGAGGCACTTTTCAGACCTGCGAATACTGGCCCGCTGGACCTCACTCAATTCGGC 1090
 QY 1091 TGCAGACTGCCCCCTGATCCATCTTTGCGGCTCCAGGCTTGGCTTTGACAGCA 1150
 Db 1091 TGCAGAACTTCCCCCGGACCCATCTTTGCGGCTCCAGGCTTCTTGTAACTGGCA 1150
 QY 1151 CACCTTTTCCCGACCTTGTCTTTAGTTTTGGGGTAAACCACTTCACTCAATTTGT 1210
 Db 1151 CGCCCTTTTCCCGACCCCTGTCTTTAGTTTTGGGGTAAACCACTTCACTCAATTTGT 1210
 QY 1211 AGCTTCTGTGGTGGGAGGCTCGAGCGGACGATGACCTGGAAACCTTGGCTCCCCAG 1270
 Db 1211 AGCTTCTGTGGTGGGAGGCTGGAGAGGATGATGACCTGGAGACTTGGATCCCCCG 1270
 QY 1271 GGGGCTCAAGGCTGCTACTTCTGGCATGCTGGTGGAGAGACTTGTGTGGAGCCG 1330
 Db 1271 GAACTTCAAGGCTGCTACTTCTGGCATGCTGGTGGAGAGACTTGTGTGGAGCCG 1330
 QY 1331 CTCTCATACCCAGCACACACAAGTGTCTGTTCTGGAGGCTGGAGGCTGGAGGCTTTA 1390
 Db 1331 CTCTCATACCCAGCACACACAAGTGTCTGTTCTGGAGGCTGGAGGCTGGAGGCTTTA 1390
 QY 1391 AGTCAAAAGCCATGGGACCCAGCCCTTATCCACTGGGTAGCCCCCGATGACCGCC 1450
 Db 1391 AGTCAAAAGCCATGGGACCCAGCCCTTATCCACTGGGTAGCCCCCGATGACCGCT 1450
 QY 1451 TGGTAGGAACTCTCAAGGACCGCTGTATGACAAATGGCAACCTTGGACATCTTCAATCA 1510
 Db 1451 TGGTAGGAACTCTTCCAGGACTGTGTATGACAAATGGCAACCTTGGACATCTTCAATTA 1510
 QY 1511 CCACATCTCAGGACAGTGTGCTTCCACTGATTTGCTGCCAAATGCTGCGGAGAGGCA 1570
 Db 1511 CCACCTCTCAGGACAGTGTGCTTCCACTGATTTGCTGCCAAATGCTGCGGAGAGGCTA 1570

QY 1571 CGGCCATGGTGGAGGTTCTCCATCTGTCAGCTGCCCAACCTCAGCAACAGCAACAGCCGCA 1630
 Db 1571 CGGCCATGGTGGAGGTTCTCCATCTGTCAGCTGCCCAACCTCAGCAACAGCAACAGCCGCA 1630
 QY 1631 CTGACACCCCAAGTCCCGCTCTCAGACATCACTGGTCTCCAGCAAGACCCCGGGAG 1690
 Db 1631 TGGCACCCCAAGTCTCGCTTTTGGACATCACAGTCTCCAGCAAGACCCCGGGAG 1690
 QY 1691 CTGAGGAGTGGGGGCGAGAGCTCCAAAAGCCCCCGGAAAGGGGTGTCTTGTGT 1750
 Db 1691 GGGGAGTGGGGCTGGGAGCCCTCCAAAAGCACCCAGAGGGCTGTCTTGTGT 1750
 QY 1751 CTGAAGTGAACCAACCTCGGCCCTGGTCAAGTGTCTGTGACAGTCCAGCACCCCGG 1810
 Db 1751 CCGATGTTACCAACCATCTGCCCTGGTCAAGTGTCTGTGACAGTCCAGCACCCCGG 1810
 QY 1811 TGAAGTGTATCAGCTGAGTACAACCTGTTCTGATGATGAGGTGCTGATCTACAGGATGA 1870
 Db 1811 TCCAGCCCTCCAAAGGCTTCTGGTCAACAACTGGTGTGACGGGACTGGCTACGACT 1930
 QY 1931 TCCAGCCCTTAACAAGGCTTCTGGTCAACAACTGGTGTGACGGGACTGGCTACGACT 1930
 QY 1931 TGTGTGCTGGCCATGTTGGGATGACAGCCACGACACTCACGGCCACCAACATGTTGG 1990
 Db 1931 TGTGGTACTGGCTATGTTGGGACGACAGCCACGCTCACAGCCACCAACATGTTGG 1990
 QY 1991 GCTGGCCAGTCTTCCACCAAGGCTACTACCCAAATGCCAGTCTATGACAGGCTCA 2050
 Db 1991 GCTGTGCCAGTCTTCCACCAAGGCTACTACCCAAATGCCAGTCTATGACAGGCTCA 2050
 QY 2051 TTTCTGGGGGACCACTGATCTTGGTCACTGGGGGCTCATCGTGGCCAGCTGGTGTCT 2110
 Db 2051 TAAAGGGGGACCACTGATCTTGGTCACTGGGAGTATCATTTGGCCACACTGCTGTCT 2110
 QY 2111 TCATCGTCACTCTATGTFGGCTACAAGGTTCGAAACAAGGCTGGGCTGGGCTGGGCTGG 2170
 Db 2111 TCATTTAATCTCATGTTGGTACAAGGTTCGAAACAAGGCTGGGCTGGGCTGGGCTGG 2170
 QY 2171 CAGCGGCC---GTGAGCAATGTTACTCGACAGCAACGGGGCCAGCCACCGCTCCAA 2227
 Db 2171 CAGCGCCACTGTACAGCAATGTTACTCGACAGCAACAAATGGGCTCAGCCCTCAGCT 2227
 QY 2228 GCAGCGCACAGCCGGGGCCCGCCGAGGGCCCGCAAGGCTGGTGGTGGCAGGAGCC 2287
 Db 2228 GCAGCGCACAGCCGGGGCCCGCCGAGGGCCCGCAAGGCTGGTGGTGGCAGGAGCC 2287
 QY 2231 GTGGATACCGTGGCCAGCTTCTCAGGACCAACCCAAAGTGTGGTGGCAATGAGT 2290
 Db 2231 GTGGATACCGTGGCCAGCTTCTCAGGACCAACCCAAAGTGTGGTGGCAATGAGT 2290
 QY 2288 TCTGGACTTCAACCGCCAGCTGGCCCGGCGCAGTGACTCTCTTCTCCAGCTCCCTGG 2347
 Db 2288 TCTGGACTTCAACCGCCAGCTGGCCCGGCGCAGTGACTCTCTTCTCCAGCTCCCTGG 2347
 QY 2348 GCAGTGGGGGCTGCGGGGCTGGGACGGGCCCCCTGGAGGATCCCACTCCGCCCCG 2407
 Db 2348 GCAGTGGGGGCTGCGGGGCTGGGACGGGCCCCCTGGAGGATCCCACTCCGCCCCG 2407
 QY 2408 GCGCCAAAGCCAGCTTCAACCGCTTGAAGGGGCTTGGGGGCTCCAGAGGCTCCCA 2410
 Db 2408 GCGCCAAAGCCAGCTTCAACCGCTTGAAGGGGCTTGGGGGCTCCAGAGGCTCCCA 2410
 QY 2411 GTCCCAAGCCAGCTTGAACCGCTTGAAGGGGCTTGGGGGCTCCAGAGGCTCC 2467
 Db 2411 GTCCCAAGCCAGCTTGAACCGCTTGAAGGGGCTTGGGGGCTCCAGAGGCTCC 2467
 QY 2468 AGAGAAAGGAGGCTGTGGACTCCAGACTCCAGCGGAGAGGGGCTGGGAGCTCGG 2527
 Db 2468 AGAGAAAGGAGGCTGTGGACTCCAGACTCCAGCGGAGAGGGGCTGGGAGCTCGG 2527
 QY 2528 CCGGGGCGCCACTCGGACCGAGCCACTGCTGGGGCCCCCTGGGGCCCCCGGGCCAGGA 2587
 Db 2528 CCGGGGCGCCACTCGGACCGGACCGACTGCTGGGGCCCCCTGGGGCCCCCGGGCCAGGA 2587
 QY 2531 CCAGAGGCGCCACTCGGACCGAGGAGGCTGCTGGGGCCCCCGGGCCAGGAGGAGG 2590
 Db 2531 CCAGAGGCGCCACTCGGACCGAGGAGGCTGCTGGGGCCCCCGGGCCAGGAGGAGG 2590
 QY 2588 GCCTCTCCCTTGGCTGGGGGCAAGGCGCAAGGCGCAAGGCGCAAGGCGCAAGGCGCA 2647
 Db 2588 GCCTCTCCCTTGGCTGGGGGCAAGGCGCAAGGCGCAAGGCGCAAGGCGCAAGGCGCA 2647
 QY 2591 GTCCTTCCCTTGGCTGGGGGCAAGGCGCAAGGCGCAAGGCGCAAGGCGCAAGGCGCA 2650
 Db 2591 GTCCTTCCCTTGGCTGGGGGCAAGGCGCAAGGCGCAAGGCGCAAGGCGCAAGGCGCA 2650
 QY 2648 ACTTTGCTGCTGGGGGCGGGAGGGGCTGTCGGGGGCGGGTACAGTCTCTCCCTGGGAGG 2707

```

2651 ACTTTCAGCTGCACTGC-----AGCTGTCCTGGTGGTACAGTCCCTTAGCCGGG 2704
2708 TCTCGAAACATCTGGACAGCGCAGCTCTCTGTCAACGGGATGCTCTTGGCCCTTTGAGG 2767
2705 TCTCAACATCTGGACAGCGCAGCTCTCTGTCAATGGCATGCTCTTGGCCCTTTGAGG 2764
2768 AGAGTGACCTGTGGGGCCGGGGACTTTGGCAGCTCCGAATGGGTGATGGAGACA 2827
2765 AGAGTGACCTGTGGGGCCGGGGACATTTGGCAGCTCAGATGGGTAAATGGAAAGTA 2824
2828 CGGTAGTGGGGTGG-GCATGCTCCCTTCTGTGGCAGGGTGGGAGAAAGGGGAAA 2886
2825 CTGTGTAGCCAGGAAGGATACCTTCCCTCCACCCCTCAAGTGGGAGAGAGATC 2884
2887 GAA--TCTCACTGGCAAGTGTGTGGAGTTCCTCAATGATGATGTTTATCATCCAGGGACAG 2944
2885 GGAGTCAGCACTGGCAAGTATTTGGAGTTCCTCAATGATGATGTTTATCATCCAGGGACAG 2944
2945 TTTGCTCTCCCTGTCAATGGCTGTGTC-----CCCCCTACCCCGCAACCCACAT 2998
2945 TCTTGTCTCCCTGTCAACAGCCTCATGCTTCTGCCCAGCTACACCCCTTCCCCTT 3004
2999 CACTTCCCA-----CCACCCGGCCGGGTGCTCAGGGAAATGGACTCGCT 3047
3005 CACTTCCAGGCTGTCCCAACCCAGCCGGGTGCTCAGGGAATG-GGACTCGCT 3063
3048 CAAATCCGGAGTGACCTGTAGTGTGGAAAGGGAGACTCCGCCCTTCTTAACACAA 3107
3064 CATGTGTCAAGTGTAGCCCTGTAGTGTGGAGAGGGCCAACTGCCCTTCTTATTACAA 3123
3108 ATCTAGCTACAGCAAGCGCTTGTGATTGCTTATG 3144
3124 ATGTAG-CGACAAGCAAGCGCTTCCGATTGCTTATG 3159

```

RESULT 2
BM787163
LOCUS K-EST0066050 S11SNU1 Homo sapiens cDNA clone S11SNU1-20-H05 5',
DEFINITION mRNA sequence.
ACCESSION BM787163 GI:19135395
VERSION 1 (bases 1 to 611)
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
Kim,Y.S.
TITLE 21C Frontier Korean EST Project 2001
JOURNAL Unpublished (2002)
COMMENT Contact: Kim YS
Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Eosaun-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongsung@mail.kribb.re.kr
Plate: 20 row: H column: 05
High quality sequence stop: 611.
Location/Qualifiers
1..611
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="S11SNU1-20-H05"
/clone_lib="S11SNU1"
/sex="M"
/tissue_type="Stomach"
/cell_type="Lymphoblast-like"
/cell_line="SNU-1"

```

/lab_host="Top10P"  

/notes:"Organ: Stomach; Vector: pME18-PL3; Site_1: XhoI;  

Site_2: XhoI; The poly (A)+ RNA was dephosphorylated with  

bacterial alkaline phosphatase (BAP) and then decapped  

with tobacco acid pyrophosphatase (TAP). The decapped  

intact mRNA was ligated with DNA-RNA linker including SfiI  

site by treatment of T4 RNA ligase and the first strand  

cDNA was synthesized with superscript II using SfiI  

oligo-dT primer. After first strand synthesis, RNA was  

degraded by NaOH treatment and cDNA was amplified by PCR  

reaction. The PCR products were digested with SfiI and  

cloned into DraIII-digested pME18S-PL3 vector. The  

obtained cDNA vectors were used for transformation of  

competent cells E. coli Top10P by electroporation method.  

The cDNA libraries constructed by this method are  

full-length enriched cDNA library."  

BASE COUNT 108 a 181 c 184 g 138 t  

ORIGIN  

Query Match 19.4%; Score 609.4; DB 14; Length 611;  

Best Local Similarity 99.8%; Pred. No. 1.2e-119;  

Matches 610; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  

OY 81 ACATTCCAAGATGCCGAGGTTCTGGGAATGCTGGGGTTCTCGATCCGGAAAATCCT 140  

DB 1 ACATTCCAAGATGCCGAGGTTCTGGGAATGCTGGGGTTCTCGATCCGGAAAATCCT 60  

OY 141 ACCGGCATCTCTTAGGGAGGATTTATTTATTTATTTTCTTTAAATCTGGAAGAAGA 200  

DB 61 ACCGGCATCTCTTAGGGAGGATTTATTTATTTATTTTCTTTAAATCTGGAAGAAGA 120  

OY 201 GAAACAAGTTGTCTTTCCCTTTCTTTCTGCTAAAGCCATGGATATAACTCAATAAG 260  

DB 121 GAAACAAGTTGTCTTTCCCTTTCTTTCTGCTAAAGCCATGGATATAACTCAATAAG 180  

OY 261 CCGCTCAGGGCTTTCCCGCGTGGACGTCGGAGCCACCATCTCCCTGCATTCGCCGGAG 320  

DB 181 CCGCTCAGGGCTTTCCCGCGTGGACGTCGGAGCCACCATCTCCCTGCATTCGCCGGAG 240  

OY 321 CCGCGGAGGGTTAGTCTGAGTCTGCTCGGGGGGGAAGGATCGCTGGCCGAGCCGGG 380  

DB 241 CCGCGGAGGGTTAGTCTGAGTCTGCTCGGGGGGGAAGGATCGCTGGCCGAGCCGGG 300  

OY 381 GAGCCCGGCGCCCGCGGAGCCGCTCGCTGGCCACCCAGCCGGGTAGATGTGCT 440  

DB 301 GAGCCCGGCGCCCGCGGAGCCGCTCGCTGGCCACCCAGCCGGGTAGATGTGCT 360  

OY 441 CGCCAGCGCTGAGTGACACAGACCATGGAGACCCCTGCTTGGTGGCCCTGCTAGCCTTGG 500  

DB 361 CGCCAGCGCTGAGTGACACAGACCATGGAGACCCCTGCTTGGTGGCCCTGCTAGCCTTGG 420  

OY 501 CATGGCGTTGGCGTGGTGGAGCGCTGCCCAAGTACTGTCTGCCAGAACTCTCTGA 560  

DB 421 CATGGCGTTGGCGTGGTGGAGCGCTGCCCAAGTACTGTCTGCCAGAACTCTCTGA 480  

OY 561 GTCACCTGGGACCTGTCCCTTCAAGGGCTGCTTTTGTACCCCTGATATTGACCG 620  

DB 481 GTCACCTGGGACCTGTCCCTTCAAGGGCTGCTTTTGTACCCCTGATATTGACCG 540  

OY 621 GCGGACAGTGGAGCTGGCCCTGGGGGCAACTTTCATATCCACATCAGCCCGCAGACTT 680  

DB 541 GCGGACAGTGGAGCTGGCCCTGGGGGCAACTTTCATATCCACATCAGCCCGCAGACTT 600  

OY 681 TGCCAAACATGA 691  

DB 601 TGCCAAACATGA 611

```

RESULT 3
BF312768
LOCUS BF312768
DEFINITION BF312768 621 bp mRNA linear EST 21-NOV-2000
601896906F1 NIH_MGC_19 Homo sapiens cDNA clone IMAGE:412624 5',
mRNA sequence.

ACCESSION BF312768
 VERSION BF312768.1 GI:11260634
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 REFERENCE 1 (bases 1 to 621)
 AUTHORS NIH-MGC http://mgi.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: c9apbs-r@mail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Ling Hong/Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
 Plate: LLCM1015 row: 1 column: 01
 High quality sequence stop: 621.

FEATURES
 source
 1..621
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:4126224"
 /clone_lib="NIH MGC 19"
 /tissue_type="neuroblastoma"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: brain; Vector: pOTB7; Site_1: XhoI; Site_2:
 EcoRI; cDNA made by oligo-dT priming. Directionally
 cloned into EcoRI/XhoI sites using the following 5'
 adaptor: GGACGAG(G). Library constructed by Ling Hong
 in the laboratory of Gerald M. Rubin (University of
 California, Berkeley) using ZAP-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies).
 Note: this is a NIH MGC Library."
 118 a 207 c 168 g 128 t

Query Match 19.3%; Score 606.8; DB 12; Length 621;
 Best Local Similarity 99.5%; Pred. No. 4.3e-119;
 Matches 619; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

806 GCCTTTGGGAGGACACCTCCGGGGCTGGTCAACTGCAGCACCCTATCGTGAACAACA 865
 Db 1 GCCTTTGGGAGGACACCTCCGGGGCTGGTCAACTGCAGCACCCTATCGTGAACAACA 60
 866 ACCAGCTGGCGGCAATCGCAGATGAGCTTTTGGAGACTTCCTGTGACATTTGGAGGATC 925
 Db 61 ACCAGCTGGCGGCAATCGCAGATGAGCTTTTGGAGACTTCCTGTGACATTTGGAGGATC 120
 926 TGGACTCTCTACAAACACTCCATGGCTCCGCTGGGACTCCGCTGGACCGCATGGTCA 985
 Db 121 TGGACTCTCTACAAACACTCCATGGCTCCGCTGGGACTCCGCTGGACCGCATGGTCA 180
 986 ACTCCACCAGCTGAGCCTGGACCAACAACCTGCTGGATCACATCGCCGAGGGCACCCTTTG 1045
 Db 181 ACTCCACCAGCTGAGCCTGGACCAACAACCTGCTGGATCACATCGCCGAGGGCACCCTTTG 240
 1046 CAGACTCGAGAAATGGCCCGCTGGATCTCACCTCCAAATCGGCTGCAAGAGCTGCCCC 1105
 Db 241 CAGACTCGAGAAATGGCCCGCTGGATCTCACCTCCAAATCGGCTGCAAGAGCTGCCCC 300
 1106 CTGATCCCAATCTTGGCCCGCTCCAGGCTTGGCTTTGACAGCCACACCCCTTTGCCCCAC 1165
 Db 301 CTGATCCCAATCTTGGCCCGCTCCAGGCTTGGCTTTGACAGCCACACCCCTTTGCCCCAC 360
 1166 CTTTGTCTTTAGTTTTGGGGTAAACCCACTTCACTGCAATTTGTAGCTTCTCTGGCTGC 1225
 Db 361 CTTTGTCTTTAGTTTTGGGGTAAACCCACTTCACTGCAATTTGTAGCTTCTCTGGCTGC 420
 1226 GGAGGCTCGAGCGGACGATGACCTGGAAAACCTGTGGCTCCCCAGGGGGCTCAAGGGTTC 1285

Db 421 GGAGGCTCGAGCGGACGATGACCTGGAAAACCTGTGGCTCCCCAGGGGGCTCAAGGGTTC 480
 1286 GCTACTTCTGGCATGTGGTGGAGGAGTGTGTGGAGCCGCTCTCATCACCAGC 1345
 Db 481 GCTACTTCTGGCATGTGGTGGAGGAGTGTGTGGAGCCGCTCTCATCACCAGC 539
 1346 ACACACAAAGTTGCTGGTTCGGAGGCCAGCGCCACACTCAAGTGCRAAAGCCATTG 1405
 Db 540 ACACACAAAGTTGCTGGTTCGGAGGCCAGCGCCAGCTCAAGTGCRAAAGCCATTG 599
 1406 GGGACCCCGAGCCCTTATCCA 1427
 Db 600 GGGACCCCGAGCCCTTATCCA 621

RESULT 4
 BQ126354
 LOCUS i116e11.y1 Melton Amplified Mouse E16 5 Pancreas 3 Mi6S1 A Mus
 DEFINITION musculus cDNA clone IMAGE:5941749 5' similar to TR:Q9ULH4 Q9ULH4
 KIAA1246 PROTEIN ; mRNA sequence.
 ACCESSION BQ126354 GI:20200265
 VERSION BQ126354.1
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 582)
 AUTHORS Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K.,
 Lemishka,T., Searce,M., Brestelli,J., Gradwohl,G., Clifton,S.,
 Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blistain,A.,
 Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J., Cardenas
 , M., Gibbons,M., McCann,R., Cole,R., Tsagarishvili,R., Williams,T.
 , Jackson,Y. and Bowers,Y.
 Endocrine Pancreas Consortium
 Unpublished (2000)
 Other ESTs: i116e11.x1
 Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
 Endocrine Pancreas Consortium
 Harvard University, Howard Hughes Medical Institute
 Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
 MA 02138
 Tel: 617-495-1812
 Fax: 617-495-8557
 Email: dmelton@biohp.harvard.edu
 Library was constructed by Dr. Douglas Melton DNA sequencing by:
 Washington University Genome Sequencing Center This clone is
 available royalty-free through LLNL; please contact the IMAGE
 consortium (info@image.llnl.gov) for further information
 Seg primer: -40RP from Gibco
 High quality sequence stop: 444.
 Location/Qualifiers
 1..582
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /clone="IMAGE:5941749"
 /clone_lib="Melton Amplified Mouse E16 5 Pancreas 3 Mi6S1
 A"
 /sex="Both"
 /tissue_type="pancreas"
 /dev_stage="Embryonic day E16.5"
 /lab_host="DH10B"
 /note="Vector: pSPOR1; Site_1: Not 1; Site_2: Sal 1;
 Library constructed using SuperScript Plasmid Library kit
 (Life Technologies). cDNA made by oligo-dT priming.
 Size-selected by column fractionation; average insert size
 0.97 kb. Amplified once on solid support. cDNA Library
 Preparation: Guolin Chen."
 126 a 194 c 158 g 104 t

FEATURES
 source
 1..582
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /clone="IMAGE:5941749"
 /clone_lib="Melton Amplified Mouse E16 5 Pancreas 3 Mi6S1
 A"
 /sex="Both"
 /tissue_type="pancreas"
 /dev_stage="Embryonic day E16.5"
 /lab_host="DH10B"
 /note="Vector: pSPOR1; Site_1: Not 1; Site_2: Sal 1;
 Library constructed using SuperScript Plasmid Library kit
 (Life Technologies). cDNA made by oligo-dT priming.
 Size-selected by column fractionation; average insert size
 0.97 kb. Amplified once on solid support. cDNA Library
 Preparation: Guolin Chen."
 BASE COUNT 126 a 194 c 158 g 104 t
 ORIGIN

Fax: 617-495-8557
 Email: dmelton@biohp.harvard.edu
 Library was constructed by Dr. Douglas Melton DNA sequencing by:
 Washington University Genome Sequencing Center This clone is
 available royalty-free through LUML; please contact the IMAGE
 consortium (info@image.llnl.gov) for further information
 Seq primer: -40UP from Gibco
 High quality sequence stop: 388.

Location/Qualifiers
 1..578
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /clone_lib="Melton Amplified Mouse E16 5 Pancreas 3 M16S1"
 A
 /sex="Both"
 /tissue_type="pancreas"
 /dev_stage="Embryonic day E16.5"
 /lab_host="DH10B"
 /notes="Vector: pSPORT1; Site_1: Not 1; Site_2: Sal 1;
 Library constructed using SuperScript plasmid library kit
 (Life Technologies). cDNA made by oligo-dT priming.
 Size-selected by column fractionation; average insert size
 0.97 Kb. Amplified once on solid support. cDNA Library
 Preparation: Guolin Chen."
 121 a 182 c 161 g 114 t

Query Match	Score	DB 14	Length	582
Best Local Similarity	100.0%	Pred. No. 8.3e-114		
Matches	582	Conservative 0	Mismatches 0	Indels 0
Gaps	0			

FEATURES	source
1408 GACCCAGCCCTTATCCACTGGGTAGCCCGGATGACCCGCTGTAGGAACTCTCA	1467
1 AGACCCAGCCCTTATCCACTGGGTAGCCCGGATGACCCGCTGTAGGAACTCTCA	60
1468 AGACCGCTGTATGACAAATGGCACTTCCATCCACACATCTCAGACAGT	1527
61 AGACCGCTGTATGACAAATGGCACTTCCATCCACACATCTCAGACAGT	120
1528 GGTGCTTCCACTGCAATCTGCTGCAATGCTGCGGAGAGCCACCGCCATGTTGAGGTC	1587
121 GGTGCTTCCACTGCAATCTGCTGCAATGCTGCGGAGAGCCACCGCCATGTTGAGGTC	180
1588 TCCATCGTCAGTGCCACACCTCAGCAACAGCACGCGCAGTCCACCCCAAGTCC	1647
181 TCCATCGTCAGTGCCACACCTCAGCAACAGCACGCGCAGTCCACCCCAAGTCC	240
1648 CGCTCTCAGACATCACTGCTCCAGCAAGACAGCCCGGGAGTGGAGCACTGGGGC	1707
241 CGCTCTCAGACATCACTGCTCCAGCAAGACAGCCCGGGAGTGGAGCACTGGGGC	300
1708 GGAGAGCTCCAAAAGCCCGGAAACGGCTGTGCTGCTGCAAGTACCACCACC	1767
301 GGAGAGCTCCAAAAGCCCGGAAACGGCTGTGCTGCTGCAAGTACCACCACC	360
1768 TCGCCCTGCTCAAGTGTCTGTCAGCAAGTCCAGCACCCCGGGTGAAGATGTACCAGTGT	1827
361 TCGCCCTGCTCAAGTGTCTGTCAGCAAGTCCAGCACCCCGGGTGAAGATGTACCAGTGT	420
1828 CAGTACAACTGCTCTGACGATGAGTACTGATTTACAGATGATCCAGCTCCACAAG	1887
421 CAGTACAACTGCTCTGACGATGAGTACTGATTTACAGATGATCCAGCTCCACAAG	480
1888 GCCTTGTGTCACAACCTGGTCTCAGGACTGGCTAGCTGTTGTGCTGGCCATG	1947
481 GCCTTGTGTCACAACCTGGTCTCAGGACTGGCTAGCTGTTGTGCTGGCCATG	540
1948 TGGGATGACAGCCAGCACACTCACGGCCACCAACATCGTG	1989
541 TGGGATGACAGCCAGCACACTCACGGCCACCAACATCGTG	582

RESULT 5
 B0126066/c
 LOCUS i11611.x1 Melton Amplified Mouse E16 5 Pancreas 3 M16S1 A Mus
 DEFINITION musculus cDNA clone IMAGE:5941749 3', similar to TR:Q9ULH4 Q9ULH4
 KIAA1246 PROTEIN i; mRNA sequence.
 ACCESSION B0126066
 VERSION B0126066.1 GI:20199977
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 1 (bases 1 to 578)
 Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K.,
 Lemishka,I., Searce,M., Brestelli,J., Gradwohl,G., Clifton,S.,
 Hillier,L., Marra,M., Pape,D., Wyllie,T., Martin,J., Blistain,A.,
 Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J., Cardenas
 ,M., Gibbons,M., McCann,R., Cole,R., Tsagarisshvili,R., Williams,T.,
 Jackson,Y. and Bowers,Y.
 Endocrine Pancreas Consortium
 Unpublished (2000)
 Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
 Endocrine Pancreas Consortium
 Harvard University, Howard Hughes Medical Institute
 Dept. of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
 MA 02138
 Tel: 617-495-1812

RESULT 6
 BM948451
 UI-M-EG0p-bvd-m-12-0-UI.r1 NIH BMAP_EG0p Mus musculus cDNA clone
 IMAGE:5690675 5', mRNA sequence.
 ACCESSION BM948451
 VERSION BM948451.1 GI:19432042
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 652)
 TITLE NIH-MGC http://mgi.nci.nih.gov/
 JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: rgs@bbs-rcmail.nih.gov
 Tissue Procurement: Dr. James Lin, University of Iowa
 cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
 This clone was contributed by the Brain Molecular Anatomy Project (BMAP)

Seq primer: pYX-5.
 Location/Qualifiers
 1..652
 /organism="Mus musculus"
 /strain="CS7BL/6"
 /db_xref="taxon:10090"
 /clone="IMAGE:5690675"
 /clone_lib="NIH BMAP_EG0p"
 /tissue_type="whole brain"
 /dev_stage="embryo 18.5 dpc"
 /lab_host="DH10B (T1 phage resistant)"
 /note="Organ: brain; Vector: pYX-Asc; Site 1: EcoR I; Site 2: Not I; The library was constructed according to Ronaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured mRNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with Not I, and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail, is CAGCCACGAC. This library was created for the University of Iowa Mouse Brain Molecular Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."
 143 a 183 c 189 g 137 t

BASE COUNT 143 a 183 c 189 g 137 t
 ORIGIN
 Query Match 16.7%; Score 525.6; DB 14; Length 652;
 Best Local Similarity 87.9%; Pred. No. 9.2e-102;
 Matches 573; Conservative 0; Mismatches 79; Indels 0; Gaps 0;
 QY 1221 GCTCGGAGGCTCGAGCGGACGATGACCTGGAAACCTGTGGCTCCCGAGGGGCTCAA 1280
 Db 1 GCTCGGAGGCTCGAGCGGATGATGACCTGGAGACCTGTGGATCCCGGAAGTCTCAA 60
 QY 1281 GGGTCGCTACTTCTGGCTGCTGGAGGAGGTTGTGCGGAGCGCCCTCTATCAC 1340
 Db 61 GGGTCGCTACTTCTGGCTGCTGGAGGAGGTTGTGCGGAGCGCCCTCTATCAC 120
 QY 1341 CCAGCACACACAAGTCTGGTTCGAGGGCCAGGGCCACACTCAAGTGGAAAGC 1400
 Db 121 CCAGCACACACAAGTCTGGTTCGAGGGCCAGGGCCACACTCTTAAGTGAAGCC 180
 QY 1401 CATTGGGGACCCAGCCGCCCTTATCCACTGGGTAGTACCCCGGATGACCGCCCTGGTAGGGAA 1460

Db 181 CATTGGGGACCCAGCCCTTTGATCCACTGGGTGGCCCGGATGACCCCTTGGTGGGAAA 240
 QY 1461 CTCCTCAAGGAGCGGTGTATGACAAATGGACCACTGGACATCTTTCATCACCACATCTCA 1520
 Db 241 CTCCTCCAGGACTGCTGTATATGACAAATGGACCACTGGACATCTTAAATACCACCTCA 300
 QY 1521 GGACAGTGGTGCCTTACCTGCATTGCTGCAATGCTCCGGAGAGGCCACCGCCATGGT 1580
 Db 301 GGACAGTGGACCCCTTTACCTGCATCGCAGCAATGACAGGAGGCTACCGGCCACTGT 360
 QY 1581 GGAGTCTCCATCTCCAGCTCCACACTCAGCAACAGCAGCCAGCCGCTGACCCCC 1640
 Db 361 AGAGTCTCCATTTGTCAGCTCCACACTTAGTAACAGCACACTAGCCGGATGGCACCCCC 420
 QY 1641 CAAGTCCCGCTCTCAGACATCACTGGCTCCAGCAAGACACCGCCGGGAGTGGAGGCGAG 1700
 Db 421 CAAGTCTCCCTTTCCGACATCAGAGTCCAGCAAGACACCGCCGGGAGGAGTAG 480
 QY 1701 TGGGGCGGAGAGCTCCCAAAAGCCCGGAAAGCGGCTGTGCTGTGTTGTTGAAGTGCAC 1760
 Db 481 TGGGGCTGGGAGCCCTCCAAAAGCACCCAGAGAGGGCTGTGCTGTGTCGGATGTTAC 540
 QY 1761 CACCACCTGGCCCTGGTCAAGTGGTCTCAGCAAGTTCAGCACCCCGGGTGAAGATGTA 1820
 Db 541 CACCACATCTCCCTGGTCAAGTGGTCTCAGCAAGTTCAGCACCCCGGGTGAAGATGTA 600
 QY 1821 CCAGCTGCAGTACAACTGCTCTGACGATGAGTACTGATTTACAGGATGATC 1872
 Db 601 TCAGCTGCAGTACAACTGCTCTGATGATGAGTACTGATTTACAGGATGATC 652

RESULT 7
 BB642310
 LOCUS BB642310
 DEFINITION BB642310 RIKEN full-length enriched, adult retina Mus musculus cDNA clone A930009011 5', mRNA sequence.
 ACCESSION BB642310
 VERSION BB642310.1 GI:15401853
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 655)
 AUTHORS Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hiramoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,J., Konno,H., Kouda,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ohno,M., Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Takeda,Y., Tanaka,T., Toya,T., Muramatsu,M. and Hayashizaki,Y.
 RIKEN Mouse ESTs (Arakawa,T., et al. 2001)
 UNPUBLISHED (2001)
 CONTACT Yoshihide Hayashizaki
 LABORATORY FOR GENOME EXPLORATION RESEARCH GROUP, RIKEN GENOMIC SCIENCES CENTER(GSC), YOKOHAMA INSTITUTE
 THE INSTITUTE OF PHYSICAL AND CHEMICAL RESEARCH (RIKEN)
 1-7-22 Suehiko-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
 TEL: 81-45-503-9222
 FAX: 81-45-503-9216
 EMAIL: genome-res@gsc.riken.go.jp/
 URL: http://genome.gsc.riken.go.jp/

Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
 Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)
 wagi,K., Fujiwaki,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuiura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
 RIKEN integrated sequence analysis (RISA) system--384-format

sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11): 1757-1771 (2000)
 Konno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara, Y. and Hayashizaki, Y.
 Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res 11 (2): 281-289 (2001)
 Yamanaka, I., Kiyosawa, H., Kondo, S., Saito, T., Shinagawa, A., Aizawa, K., Fukuda, S., Hara, A., Itoh, M., Kawai, J., Shibata, K., Arakawa, T., Ishii, Y. and Hayashizaki, Y.
 Mapping of 19032 mouse cDNAs on mouse chromosomes. J. Struct. Func. Genomics 2 pre, L72-L86 (2001)
 Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.
 e mouse tissues.

FEATURES source

1. .655
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /clone_lib="RIKEN full-length enriched, adult retina"
 /clone_lib="RIKEN full-length enriched, adult retina"
 /tissue_type="retina"
 /dev_stage="adult"
 /lab_host="DH10B"
 /note="Site 1: Sali; Site 2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5' GAGAGAGAGAGATCCAGAGCTCTTTTTTTTTTTTNN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 20.0 and subtraction to Rot = 459.0. Second strand cDNA was prepared with the primer adapter of sequence [5'GAGAGAGAGATTCGATTAATAAATATATCCCGCCCGCC 3']. cDNA was cleaved with XhoI and BamHI. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLC I. -Retina RNA was provided by Stefano Gustincich, Department of Neurobiology, Harvard Medical School, 220 Longwood Ave., Boston, MA02115, USA, whose assistance we gratefully acknowledge."
 119 a 203 c 188 g 144 t 1 others

Query Match 16.4%; Score 514.2; DB 10; Length 655;
 Best Local Similarity 87.8%; Pred. No. 2.5e-99;
 Matches 574; Conservative 0; Mismatches 74; Indels 6; Gaps 1;
 335 AGCTCGAGTGTCTCGGGGGGAGAGTGGTGGCGGAGCCGGGGCGGCC 394
 2 AGCTCGAGTGTCTCGGGGGGAGAGTGGTGGCGGAGCCGGGGCGGCC 61
 395 CGCGAGCGGCTCGGTGCCACCCAGCGGGGGTGTGCTGCTCGCCAGG----- 448
 62 TCGCGAGCGGCTGTGAGCTCTCCCGCCACGCGGTAGATGCCCGCTTCCCGAGTGCAGA 121
 449 CGCTGAGTGCAGACCATGGAGCCCTGCTGGTGGCTGTAGGTTTGGCATGGCGT 508
 122 GGCTGAGTGCAGACCATGGAGACTGCTGGTGGCTGTGGCTTTTGGCATGGCGT 181
 509 TTGCGGTGGTGCAGCGCTGCCCAAGTACTGTGTCTGCCAGAAATCTGTCTGAGTCACTGG 568
 182 TTGCTGTGGTGCAGTCTGCGCCCAAGTACTGCGTCTGCCAGAAATCTGTCTGAGTCACTGG 241
 569 GGACCCCTGTGCCCTTCCAAAGGGGTGCTTTGTACCCCTGTATTTGACCGCGGACAG 628
 242 GGACCCCTGTGCTCCCAAGGGGTGCTTTGTGCCCCCTGACATTTGACCGGAGACGG 301
 629 TGGAGCTGGCTGGCGGGCAACTTCATCTCCACATCCAGCCGACGACTTTGCCAACA 688

BASE COUNT ORIGIN

302 TGGAGCTGGCTCTGGGTGGCAACTTCATCATCCATTCACATTTGGCCGCCAAGACTTTGGCCAACA 361
 689 TGACGGGGCTGGTGGACCTGACCTCCAGGAAACACCATCAGCCACATCCAGCCCTTTT 748
 362 TGACAGGGCTGGTGGATCTGACCTTGTCCAGGAACACCATCAGCCATATCCAGCCCTTCT 421
 749 CCTTTCTGGACCTCGAGAGCTCCGGTCCCTCGCATTTGACAGCAATCGGCTGCCAAGCC 808
 422 CCTTCTGGACCTTGGAGGCTCCGGTCCCTGACAGCAATCGGCTGCCAAGCC 481
 809 TTGGGGAGGACACCTCCGGGGCTGGTCAACCTTGGAGCACTTATCGTCAAAACAACAAC 868
 482 TTGGGGAGGACACACTCCGGGGTCTGGTCAACCTTGGAGCACTTATTTGTGACAAATAACC 541
 869 AGCTGGGGGATCGAGATGAGGCTTTTGGAGACTTCTGTCGACATTTGGAGGATCTGG 928
 542 AGCTGGGGGATCGAGATGAGGCTTTTGGAGACTTCTGTCGACATTTGGAGGATCTGG 601
 929 ACCTCTCTTACAAACACTTCCATGCTGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGG 982
 602 ACCTATCTTACAAACACTTCCAGGACTTCCCTGCTGGATTCCTGACCTTTTGAAGACCTT 655

RESULT 8 BC007718

LOCUS BC007718 1853 bp mRNA linear HTC 12-JUL-2001
 DEFINITION Homo sapiens, hypothetical protein MGC3103, clone IMAGE:4110922, mRNA.
 ACCESSION BC007718 GI:14712783
 VERSION BC007718.1
 KEYWORDS HTC.
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 1853)
 AUTHORS Strausberg, R.
 DIRECT SUBMISSION
 TITLE Submitted (11-MAY-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
 REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 COMMENT Contact: MGC help desk
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC),
 Gaithersburg, Maryland;
 Web site: <http://www.nisc.nih.gov/>
 Contact: nisc.mgc@nhgri.nih.gov
 Shenchenko, Y., Wetherby, K.D., Beckstrom-Sternberg, S.M., Benjamen, B., Blakesley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S., Dietrich, N.L., Guan, X., Gupta, J., Ho, S.-L., Karlins, E., Legaspi, R., Lim, M., Maduro, Q.L., Masello, C., Mastrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Snyder, B., Stantripop, S., Thomas, P.J., Tjongson, E.E., Touchman, J.W., Tsurgeon, C., Vogt, J.L., Walker, M.A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAL Plate: 17 Row: e Column: 2
 This clone was selected for full length sequencing because it passed the following selection criteria: Genomescan gene prediction
 This clone has the following problem: incomplete processing.

FEATURES source

1. .1853
 /organism="Homo sapiens"
 /db_xref="locusID:78998"
 /db_xref="taxon:9606"
 /clone="IMAGE:4110922"

/tissue_type="Muscle, rhabdomyosarcoma"
/clone_lib="NIH_MGC_17"
/lab_host="DH10B-R"
/note="Vector: pOTB7"
BASE COUNT 329 a 648 c 558 g 318 t
ORIGIN
Query Match 15.7%; Score 492.6; DB 11; Length 1853;
Best Local Similarity 63.2%; Pred. No. 1.1e-94;
Matches 828; Conservative 0; Mismatches 444; Indels 39; Gaps 3;

QY 520 GACGCTGCCCCAAGTACTGCTGCGCAACATCTGCTGCTGAGTCACTGGGGACCCTGTGC 579
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 272 GCCCGCTGCCGCTGCGCTGCGCAGAACCTGTCCGAGTGCCTCAGCACCCCTGTGT 331
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 580 CCTTCAAGGGTGTCTTTGTACCCCTGTATTTGACCGCGGACAGTGGAGTGCCTC 639
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 332 GCCACAGGGTCTGCTTTGTGCGGCCCAAGTGGACCGGGCCACATGAGTGTGG 391
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 640 CTGGGGCAATTATACATACATACATACATACATGCGCCAGGACTTTGGCCAAATGACGGGGTG 699
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 392 CTGGGTGCAAACTTATCAGGCCCCCTGGGGCCCCCTGACTTCGGCAACATGACGGGACTG 451
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 700 GTGGACCTGACCTGCTCCAGAACACCATCAGCCACTCAGCCCTTTTTCCTTTCGAC 759
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 452 GTGGACCTGACACTGTCTGCAATGCTGCAATGCTGCAATGCTGCAATGCTGCAATGCTGCAAT 511
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 760 CTCGAGAGCTCCGCTCCCTGTGATCTTGACAGCAATGCGCTGCAAGCCCTTGGGGAGAC 819
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 512 CTCGAGAGCTCCGCTCCCTGTGATCTTGACAGCAATGCGCTGCAAGCCCTTGGGGAGAC 571
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 820 ACCCTCGGGGCTGGTCAACCTGACGACCTTATCTGTGAACAAACCAACCAACGCTGGGCGGC 879
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 572 AGCTCGGGGCCCCGTCAATCTGACGACCTATCTCAGCGGCAACCAACGCTGGGCGGC 631
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 880 ATCCGAGATGAGCTTTTGGAGACTTCTGCTGATCTGAGAGTCTGGAGTCTGGACCTTCTTAC 939
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 632 ATCCGCGGGGAGCTTTCAGCAGCTTCTAGAGGCTTGGAGGACCTGAGACCTGTCCTAC 691
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 940 AACAACCTCAATGCTGCCTGGGACTCCGTGGACGATGCTGCAACCTTCCACCACTCCACCACTG 999
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 692 AACAACCTCGGAGGCTTCCGCGGACTTCCAGGAGGCTTCCAGGAGGCTTCCAGGAGGCTTCC 751
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 1000 AGCTGGACCAACCTGTGTGATCAATCGCCGAGGGGACCTTTGAGACACTTCAGAA 1059
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 752 AACCTGGACATAACCTTATGAGGACTGCCCCAGGCGCTTTGGCCAGCTCGGTAC 811
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 1060 CTGGCCGCTGGATCTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTTT 1119
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 812 CTCTCCCGCTGGACCTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTTTTC 871
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 1120 GCCCGCTCCAGCTTCGGCTTTGACAGCCACACCCTTTGCCACCCTTGTCTTTAGT 1179
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 872 TCTGTGGCGTGTAT-----GCAGAGGCTCTCCCGCCCTGTGGCTGTCGAGC 919
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 1180 TTTGGGGTAACCCACTTCACTGCAATTTGAGTCTTCTGGCTGGAGGCTCGAGCGG 1239
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 920 TTTAGGGAAACCCCTGCTCACTGCACTGCACTGCACTGCTGCTGGCTGGGCGCTGGGCGG 979
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 1240 GAGATGACCTGGAACCTGTGGCTTCCAGGGGGCTCAAGGTGGCTACTTTGGGAT 1299
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 980 CCGAGGACTGGAAAACGTGGCCTCCCGCGCGGCTGCGGGCGCTACTTTCTGGGCA 1039
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 1300 GTGGCTGAGAGAGTTTGTGGAGCGCTCATCATCCAGCACACACACACAGTTC 1359
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1040 GTGCCGGAGGGGAGTTCCTCTGTGAGCGGCCCTCATTTGCGGCGCCACACAGGCGCTC 1099
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 1360 CTGGTTCTGGAGGGCCAGGGGCGCCACTCAAGTGCAGAACCAATGGGACCCCAAGCC 1419
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1100 TGGTGTGTAAGGCCAGCGGGCCAGCTGCTGGCTGGCGGCCCTGGGTGACCCCGGCCCT 1159
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 1420 CTTATCCACTGGTACCCCGGAGTATGACCGCTGGTTAGGAACTCTCTCAAGGACCGGTGTC 1479
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

1160 ACCATGCACTGGGTGGCTTCGAGCACCGGTGGTTGGCAACTCTCCCGAGCCCGGGCT 1219
1480 TATGCAATGGCACCCCTGGACATCTTATCACCAATCTCAGGACAGTGGTCCCTTCCAC 1539
1220 TTCCCAACGGGACCTTATAGATTTGGGGTGCACCGGCTGGGGGACCTGCGGCTACACC 1279
1540 TGCAATGCTGCCAATGCTGCGGAGAGCCACGGCCATGCTGGAGGCTTCCATCGTCCAG 1599
1280 TGCATCGCCACCAACTGCTGTGTGGAGCCACAGCCGAGTAACTTGGGGTGTGGCC 1339
1600 CTGGCCACACCTCAGCAACAGCAGCCGCACTGCACCCCAAGTCCCGCCCTCTCAGAC 1659
1340 TTGGCCCAATGTTGGGAACAGAGTCCCGAGGGGGGCCGCC-----CGGGCCCTTCGAC 1393
1660 ATCACTGCTCCAGCAAGACCAAGCCGGGGAGTGGAGGAGTGGGGGGGAGAGGCTCC 1719
1394 ATCGCGGCTCCGCTCGCACCTCTCCAGGAGTGGGGGACCTGGAGTCTGAGCC---- 1449
1720 AAAAAACCCCGGAACGGGCTGCTTTGTGTGAAGTGAACCAACCACTCGGCCCTGGTFC 1779
1450 -----AGCCCTGCAAGTGAAGGAGTACCGCCACCTCAGGGGCTGTG 1492
1780 AAGTGGTCTGTCAGCAAGTCAAGCACCAGCGGTGAAGTGTACCAGCTGCAG 1830
1493 AGCTGGGTTCGCGGGCCAGCCACCAGTGTGATGTTCCAATCCAG 1543

RESULT 9
AW163387
LOCUS au94f03.y1 Schneider fetal brain 00004 Homo sapiens cDNA clone
DEFINITION IMAGE:2783933 5', mRNA sequence.
ACCESSION AW163387
VERSION AW163387.1 GI:6302420
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 457)
AUTHORS Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
Krisman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin
White,Y., Wyllie,T., Scheinberg,K., Steptoe,M., Tan,F., Theising,B.,
J., Moore,B., Schellenger,K., Waterston,R. and Wilson,R.
WASHU-NCI human EST Project
TITLE Unpublished (1997)
JOURNAL
COMMENT Other_ESTs: au94f03.x1
Contact: wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LILNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -40RP from Gibco
High quality sequence stop: 435.
Location/Qualifiers

FEATURES
source
1..457
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2783933"
/clone_lib="Schneider fetal brain 00004"
/sex="male"
/tissue_type="frontal lobe"
/dev_stage="5 months post-conception"
/lab_host="DH10B"
/note="Organ: brain; Vector: pBluescript SK (Stratagene);
Site 1: SstI; Site 2: XhoI; Double-stranded cDNA was
prepared from human fetal brain tissue. 5' and 3'
adaptors were used in cloning as follows: 5' adaptor
sequence:
5'-GAGAGAGAGAGAGACTCAAGGATCCTTAATTAATTAATCCCTCCCTCCCCCCCC-3'

Email: est@watson.wustl.edu
 This clone is available royalty-free through LILNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 Seg primer: -4ORP from Gibco
 High quality sequence stop: 431.
 Location/Qualifiers

FEATURES

1..435
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="Schneider fetal brain 00004"
 /sex="male"
 /tissue_type="frontal lobe"
 /dev_stage="5 months post-conception"
 /lab_host="DH10B"
 /note="Organ: brain; Vector: pBluescript SK (Stratagene);
 Site 1: SstI; Site 2: XhoI; Double-stranded cDNA was
 prepared from human fetal brain tissue. 5' and 3'
 adaptors were used in cloning as follows: 5' adaptor
 sequence:
 5'-GAGAGAGAGAGAGCTCAAGGATCCCTTAATTAATAATCCCCCCC-3'
 and 3' adaptor sequence:
 5'-GAGAGAGAGACTCGAGTTTTTTTTTTTTTTTTTT-3'. The library was
 size-selected for >0.5 kb inserts and has an average
 insert size estimated at 1.2 kb. This library was
 constructed using the CAP-trapper method for full-length
 enrichment and has not undergone amplification. Library
 was constructed by Dr. Claudio Schneider (LNCIB-Area
 Science Park, Trieste, Italy). 1 others
 80 a 126 c 131 g . 97 t

BASE COUNT

ORIGIN

Query Match 13.4%; Score 422.8; DB 10; Length 435;
 Best Local Similarity 98.2%; Pred. No. 7, 5e-80;
 Matches 427; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

66 ACTTGAAGACACCCCAATTCCTCAAGATGCGCGAGGTTCTGGGAATGCTGGGTTCTTC 125
 1 ACTTGAAGACACCCCAATTCCTCAAGATGCGCGAGGTTCTGGGAATGCTGGGTTCTTC 60
 126 GATCGGAAATCTCCTCCGATCCTCTAGGAGGATTAATTAATTAATTTCTTTA 185
 61 GATCGGAAATCTCCTCCGATCCTCTAGGAGGATTAATTAATTAATTTCTTTA 120
 186 ATCTGGAAGAGAGAAACAAGTTGTCTTTTCCCTTCTTCTTAAAGCCATGG 245
 121 ATCTGGAAGAGAGAAACAAGTTGTCTTTTCCCTTCTTCTTAAAGCCATGG 180
 246 ATATAACTGAATAAGCGCTCAGGCTTTCCCGCTGGACGTCGAGGCCACCACTGC 305
 181 ATATAACTGAATAAGCGCTCAGGCTTTCCCGCTGGACGTCGAGGCCACCACTGC 240
 306 CTGCATTCGCGGAGCGCGCGGAGGTTTAGCTCGAGTCTGTCTCGGCGGGAGGATG 365
 241 CTGCATTCGCGGAGCGCGCGGAGGTTTAGCTCGAGTCTGTCTCGGCGGGAGGATG 300
 366 CGTGGCCGAGCGGGGAGCCCGGCGCCCGGAGCGGCTCGGTGCCACCCAGCCGG 425
 301 CGTGGCCGAGCGGGGAGCCCGGCGCCCGGAGCGGCTCGGTGCCACCCAGCCGG 360
 426 GGGTAGATGCTGCTCGCCGAGCGCTGAGTGACCAACCATGGAGCCCTCTTTGGTG 485
 361 GGGTAGATGCTGCTCGCCGAGCGCTGAGTGACCAACCATGGAGCCCTCTTTGGTG 420
 486 CCTCCTAGCTTTGG 500
 421 CCTCCTAGCTTTGG 435

RESULT 11
 BB278209
 LOCUS
 DEFINITION
 685 bp mRNA linear EST 31-AUG-2001
 RIKEN full-length enriched, adult retina Mus musculus cdNA

and 3' adaptor sequence:
 5'-GAGAGAGACTCGAGTTTTTTTTTTTTTTTTTT-3'. The library was
 size-selected for >0.5 kb inserts and has an average
 insert size estimated at 1.2 kb. This library was
 constructed using the CAP-trapper method for full-length
 enrichment and has not undergone amplification. Library
 was constructed by Dr. Claudio Schneider (LNCIB-Area
 Science Park, Trieste, Italy). 1 others
 83 a 131 c 140 g 103 t

BASE COUNT

ORIGIN

Query Match 14.1%; Score 441.8; DB 10; Length 457;
 Best Local Similarity 99.3%; Pred. No. 6, 7e-84;
 Matches 454; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

66 ACTTGAAGACACCCCAATTCCTCAAGATGCGCGAGGTTCTGGGAATGCTGGGTTCTTC 125
 1 ACTTGAAGACACCCCAATTCCTCAAGATGCGCGAGGTTCTGGGAATGCTGGGTTCTTC 60
 126 GATCGGAAATCTCCTCCGATCCTCTAGGAGGATTAATTAATTAATTTCTTTA 185
 61 GATCGGAAATCTCCTCCGATCCTCTAGGAGGATTAATTAATTAATTTCTTTA 120
 186 ATCTGGAAGAGAGAAACAAGTTGTCTTTTCCCTTCTTCTTAAAGCCATGG 245
 121 ATCTGGAAGAGAGAAACAAGTTGTCTTTTCCCTTCTTCTTAAAGCCATGG 180
 246 ATATAACTGAAT-AAGCGCTCAGGCTTTCCCGCTGGACGTCGAGGCCACCACTGC 304
 181 ATATAACTGAAT-AAGCGCTCAGGCTTTCCCGCTGGACGTCGAGGCCACCACTGC 240
 305 CCGTCAATTCGCGGAGCGCGCGGAGGTTTAGCTCGAGTCTGTCTCGGCGGGAGGAT 364
 241 CCGTCAATTCGCGGAGCGCGCGGAGGTTTAGCTCGAGTCTGTCTCGGCGGGAGGAT 300
 365 GGTGGCCGAGCGGGGAGCCCGGCGCCCGGAGCGGCTCGGTGCCACCCAGCCGG 424
 301 GGTGGCCGAGCGGGGAGCCCGGCGCCCGGAGCGGCTCGGTGCCACCCAGCCGG 360
 425 GGGTAGATGCTGCTCGCCGAGCGCTGAGTGACCAACCATGGAGCCCTCTTTGGTG 484
 361 GGGTAGATGCTGCTCGCCGAGCGCTGAGTGACCAACCATGGAGCCCTCTTTGGTG 420
 485 GCCTGCTAGCTTTGGGATGGCTTTCCCTGCTCGA 521
 421 GCCTGCTAGCTTTGGGATGGCTTTCCCTGCTCGA 457

RESULT 10
 AW161566
 LOCUS
 DEFINITION
 IMAGE:2781720 5', mRNA sequence.
 AW161566
 AW161566.1 GI:6300599
 EST.
 HUMAN.
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 435)
 Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S.,
 Krieman, D., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marr, M., Martin
 , J., Moore, B., Schellenberg, K., Steptoe, M., Tan, F., Theising, B.,
 White, Y., Wylie, T., Waterston, R. and Wilson, R.
 WashU-NCI Human EST Project
 Unpublished (1997)
 Other ESTs: au71e01.x1
 Contact: Wilson RK
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810

clone A930009011 3', mRNA sequence.
 BB278209
 VERSION BB278209.2 GI:15410982
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. (bases 1 to 685)
 AUTHORS Arakawa, T., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hiramoto, K., Hori, F., Ishii, Y., Ito, M., Kawai, J., Konno, H., Kouda, M., Koya, S., Matsuyama, T., Miyazaki, A., Nomura, K., Ohno, M., Okada, Okazaki, Y., Okido, T., Saito, R., Sakai, C., Sakai, K., Sano, M., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagami, M., Tegawa, A., Takahashi, F., Takeda, Y., Tanaka, T., Toya, T., Muramatsu, M. and Hayashizaki, Y.

TITLE RIKEN Mouse ESTs (Arakawa, T., et al. 2001)
 JOURNAL Unpublished (2001)
 COMMENT On Jul 7, 2000 this sequence version replaced gi:8975231.
 Contact: Yoshihide Hayashizaki
 Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute
 The Institute of Physical and Chemical Research (RIKEN)
 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
 Tel: 81-45-503-9222
 Fax: 81-45-503-9216
 Email: genome-res@gsc.riken.go.jp,
 URL: http://genome.gsc.riken.go.jp/
 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.

Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)
 wagi, K., Fujiwaki, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.

RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)
 Konno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara, Y. and Hayashizaki, Y.
 Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
 Yamanaka, I., Kiyosawa, H., Kondo, S., Saito, T., Shinagawa, A., Aizawa, K., Fukuda, S., Hara, A., Itoh, M., Kawai, J., Shibata, K., Arakawa, T., Ishii, Y. and Hayashizaki, Y.
 Mapping of 19032 mouse cDNAs on mouse chromosomes. J. Struct. Func. Genomics 2 pre, L72-L86 (2001)
 Please visit our web site (<http://genome.gsc.riken.go.jp/>) for further details.

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN, Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

FEATURES
 source
 1..685
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /clone="A930009011"
 /clone_lib="RIKEN full-length enriched, adult retina"
 /tissue_type="retina"
 /dev_stage="adult"
 /lab_host="DH10B"
 /note="Site 1: SalI; Site 2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer {5'

GAGAGAGAGAGAGATCCAGAGCTCTTTTTTTTTTTTTTTTTTTT 3', cDNA was prepared by using trihalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 20.0 and subtraction to Rot = 459.0. Second strand cDNA was prepared with the primer adapter of sequence 5'GAGAGAGAGATTCGAGTTAAATAAATACCCCTCCCTCC 3'. cDNA was cleaved with XhoI and BamHI. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLC I. -Retina RNA was provided by Stefano Gustinich, Department of Neurobiology, Harvard Medical School, 220 Longwood Ave., Boston, MA02115, USA, whose assistance we gratefully acknowledge."
 BASE COUNT 146 a 199 c 196 g 144 t
 ORIGIN

Query Match 12.0%; Score 377.2; DB 10; Length 685;
 Best Local Similarity 78.9%; Pred. No. 4.3e-70;
 Matches 546; Conservative 0; Mismatches 118; Indels 28; Gaps 7;
 QY 2473 AAGGAGAGCTGCTGGACTCCAGGACTCCAGCCGGAGAGGGCTGGGACGTCGCCCCGG 2532
 Db 1 AAAGAAGAACTTCTAGACTCCAGGACTCCAGCAGGAGAGGGCCAGGACATCATCCAGA 60
 QY 2533 GCCACCACTGGACCAGAGCCACTGCTGGGCCCCCTGGGCCCCGGCCAGGACGCTG 2592
 Db 61 GGGCCACCACTCAGACCAGAGCCACTGCTGGGCCCCGGCCAGGACGCTGTT 120
 QY 2593 CTCCTCCCTGGCTGGAGGGAAGCCAAACGACGACTCCTCGACATGGGGACTTTT 2652
 Db 121 CTCCTCCCTGGCTGGAGGGAAGCCAAACGACGACTCCTCGACATGGGGACTTTT 180
 QY 2653 GCTGTGGGGGGGGGGGGGCTGCTGGGGGGGGCTACAGTCTCCCTCGGAAGTCTCG 2712
 Db 181 GCAGCTGGAGCTGC-----AGCTGTCTCTGGTGTAGACTGCTCCCTTTGAGGAGAGT 234
 QY 2713 AACATCTGGACGAAGCGGAGCCCTCTGTAAACGGGATGCTCTTGGCCCTTTGAGGAGAGT 2772
 Db 235 AACATCTGGACGAAGCGGAGCCCTCTGTAAACGGGATGCTCTTGGCCCTTTGAGGAGAGT 294
 QY 2773 GACCTGGTGGGGGGGGGGGACTTTGGCAGCTCGGAATGGGTGATGGAGAGCAGGTC 2832
 Db 295 GACCTGGTGGGGGGGGGGGACTTTGGCAGCTCGGAATGGGTGATGGGAAGTACTGTG 354
 QY 2833 TAGTGTGGGGGTGG--GCATGCTCCCTTTCTGTGCGCAGGGTGGGAGAAGGGAAAGAA-- 2889
 Db 355 TAGCCAGGAGAGGAGTACCTCCCTCCACCCCTCAAAGTGGGAGAGGAGTGGAGT 414
 QY 2890 TCTACTGGCAAGTGTGTTGGAGTTCATGGTGTATGTTTACATCCAGGGACAGTTCG 2949
 Db 415 CAGCACTGGCAAGTGTGTTGGAGTTCATGGTGTATGTTTACATCCAGGGACAGTTCG 474
 QY 2950 TCTCCCTGTCAATGGCTCGTGC-----CCCCCTACCCCGCAACACCCACATCACCT 3003
 Db 475 TCTCCCTGTCAACAGCCCTCAATGCTCTTGTCCAGCTACACCCCTTCCCCATGTACCT 534
 QY 3004 CCCCC-----CCACCCCGGGGGTGTGTCTCAGGGAATGTGCACTCGCTCAAAAT 3052
 Db 535 TCCAGGCTGTCCCCCACCCAGGGGGTGTGTCTCAGGGAATG--GGACTGCCTCATGT 593
 QY 3053 GCCGACTGAGCCCTGAGTGTGTTGAAAGGGGAGACTCGCCCTTCTTAATCAAAATGTA 3112
 Db 594 GTCAGACTGAGCCCTGAGTGTGTTGAGAGGCCAACAACCTGCTTCTATTCAAAATGTA 653
 QY 3113 GCCTACAAGCAGCGGCTTTGGATTGCTTATG 3144
 Db 654 G-CGAAAGCAAGCGGCTTCGGATTGCTTATG 684

RESULT 12
 BML193100
 LOCUS BML193100 357 bp mRNA linear EST j3-DEC-2001
 DEFINITION TCBAP1D9544 Pediatric pre-B cell acute lymphoblastic leukemia

Db 241 TATGCAATGGCCCTGGACATCTTTCACACACATCTCAGGACAGTGGTCCCTCACC 300
 BG380772/c
 QY 1540 TGCATTGCTGCCAATGCTCCCGAGAGCCCGGCGCATGGTGGAGGTCTCCAT 1592
 Db 301 TGCATTGCTGCCAATGCTCCCGAGAGCCCGGCGCATGGTGGAGGTCTCAT 353

RESULT 13
 BG380772/c
 LOCUS 635 bp mRNA linear EST 12-MAR-2001
 DEFINITION UI-R-CTO-bub-c-04-0-UI.61 UI-R-CTO Rattus norvegicus cDNA clone
 UI-R-CTO-bub-c-04-0-UI 3', mRNA sequence.
 ACCESSION EG380772
 VERSION EG380772.1 GI:13305244
 KEYWORDS EST.
 SOURCE Norway rat.
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus
 REFERENCE 1 (bases 1 to 635)
 AUTHORS Bonaudo, M.F., Lennon, G. and Soares, M.B.
 TITLE Normalization and subtraction: two approaches to facilitate gene
 discovery
 JOURNAL Genome Res. 6 (9), 791-806 (1996)
 MEDLINE 97044477
 COMMENT Contact: Soares, MB
 Program for Rat Gene Discovery and Mapping
 University of Iowa
 451 Eckstein Medical Research Building Iowa City, IA 52242, USA
 Tel: 319 335 8250
 Fax: 319 335 9565
 Email: msoares@blue.weeg.uiowa.edu
 The sequence contained an oligo-dT track that was present in the
 oligonucleotide that was used to prime the synthesis of first
 strand cDNA and therefore this may represent a bonafide poly A
 tail. The sequence tag present in the cDNA between the NotI site
 and the oligo-dT track served to verify it as a clone from the
 Soares Lab Clone distribution. Clones will be available through
 Research Genetics (www.resgen.com)
 Seq primer: M13 Forward
 POLYA=Yes.
 Location/Qualifiers
 1..635
 /organism="Rattus norvegicus"
 /strain="Sprague-Dawley"
 /db_xref="taxon:10116"
 /clone="UI-R-CTO-bub-c-04-0-UI"
 /dev_host="UI-R-CTO"
 /dev_stage="ADULT"
 /note="Vector: pT73D-pac (Pharmacia) with a modified
 polylinker; Site 1: Not I; Site 2: Eco RI; The UI-R-CTO
 library is a normalized library constructed from the
 following rat brain tissues: embryonic day 17, embryonic
 day 19, embryonic day 21, adult day 1, adult day 12, adult
 day 75, adult day 200. For a detailed description of the
 library from which this clone was derived, please visit
 our web site at ratseq.eng.uiowa.edu. The subtraction has
 been previously described in (Bonaudo, Lennon and Soares,
 Genome Research 6:791-806, 1996)
 TAG LIB=UI-R-CTO
 TAG TISSUE=rat brain pool
 TAG SEQ=ACTTC"
 135 a 174 c 183 g 143 t

BASE COUNT 135 a 174 c 183 g 143 t
 ORIGIN
 Query Match 10.4%; Score 327; DB 12; Length 635;
 Best Local Similarity 77.4%; Pred. No. 2, 2e-59;
 Matches 485; Conservative 0; Mismatches 110; Indels 32; Gaps 6;

Baylor-HGSC project=TCBA Homo sapiens cDNA clone TCBA9544, mRNA
 sequence.
 BM193100
 BM193100.1 GI:17651296
 EST.
 human.
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 357)
 AUTHORS Wei, Y., Tsang, Y.T.M., Mei, G., Ku, J.M., Ali-Osman, F.R. Jr.,
 Gunaratne, P.H., Muzny, D., Bouck, J., Gibbs, R.A. and Margolin, J.F.
 Pediatric Leukemia cDNA Sequencing Project (2001)
 TITLE Pediatric Leukemia cDNA Sequencing Project (2001)
 JOURNAL Unpublished (2001)
 COMMENT Contact: Dr. Judith F. Margolin
 Texas Children's Cancer Center and Human Genome Sequencing Center
 at Baylor College of Medicine
 1102 Bates, MC3-3320 Houston, TX 77030, USA
 Tel: 832-824-4536
 Fax: 832-825-4038
 Email: clones@txccc.org
 Seq primer: M13 primer.
 Location/Qualifiers
 1..357
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="TCBA9544"
 /clone_lib="pediatric pre-B cell acute lymphoblastic
 leukemia Baylor-HGSC project=TCBA"
 /sex="male"
 /tissue_type="leukopheresis"
 /cell_type="pre-B cell"
 /dev_stage="pediatric 2 years"
 /lab_host="DH10B"
 /not_e="vector; lambda PSB; Site 1: BamHI; Site 2: EcoRI;
 First strand cDNA was primed with an anchored
 3'; V=A,C,G; N=A,C,G,T) and then dG tailed. Second strand
 was primed with a BamHI-dC primer
 [5'AGAGGCTCGATCGCGCGCGCAATAATAAT(C) 3'].
 Double-stranded cDNA was then digested with BamHI and XhoI
 and directionally cloned into the BamHI and Sali sites of
 lambda PSB vector. Library was through one round of
 normalization. Library was constructed by Wei Yu at RIKEN
 of Japan (Carninci P, Westover A, Nishiyama Y, Ohsumi T,
 Itoh M, Nagaoka S, Sasaki, Okazaki Y, Muramatsu M,
 Schneider C, Hayashizaki Y, High efficiency selection of
 full-length cDNA by improved biotinylated cap trapper...
 DNA Res 4: 1, 61-6, Feb 28 1997)"
 71 a 111 c 103 g 72 t

BASE COUNT 71 a 111 c 103 g 72 t
 ORIGIN
 Query Match 11.0%; Score 346.6; DB 13; Length 357;
 Best Local Similarity 98.9%; Pred. No. 1.3e-63;
 Matches 349; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 QY 1240 GACGATGACTGGAACCTGTGGCTCCCGAGGGCCCTCAAGGTCGGTACTTCTGGCAT 1299
 Db 1 GACGATGACTGGAACCTGTGGCTCCCGAGGGCCCTCAAGGTCGGTACTTCTGGCAT 60
 QY 1300 GTCGTTGAGGAGGTTGTGTGGAGCGCCCTCTCATCCAGCACACACAAGTTG 1359
 Db 61 GTCGTTGAGGAGGTTGTGTGGAGCGCCCTCTCATCCAGCACACACAAGTTG 120
 QY 1360 CTGGTTCTGAGGGCCAGGGCCACACTCAAGTGCAGAACCCATTGGGACCCAGCCCC 1419
 Db 121 CTGGTTCTGAGGGCCAGGGCCACACTCAAGTGCAGAACCCATTGGGACCCAGCCCC 180
 QY 1420 CTTATCCACTGGTGTAGCCCCGATGACCGCTGGTAGGGAATCTCTCAAGGACCCGCTGTC 1479
 Db 181 CTTATCCACTGGTGTAGCCCCGATGACCGCTGGTAGGGAATCTCTCAAGGACCCGCTGTC 240
 QY 1480 TATGACAATGGCCCTGGACATCTTTCACACACATCTCAGGACAGTGGTCCCTCACC 1539

QY 2533 GGCACCACTCGGACCGAGCCACTGTGGGGCCCTTGGGGCCCGCCGAGGACCCCTG 2592
 |||||
 Db 635 GGCACCACTCAGACCGAGCCACTGTGGGGCCCGCCGAGGACCCCTT 576
 |||||
 QY 2593 CTCCTTCCTGGGCGGAAAGGCGAAACAGCCACTCTTGGACATGGGACATTT 2652
 |||||
 Db 575 CTCCTTCCTGGGCGGAAAGGCGAAACAGCCACTCTTGGACATGGGACATTT 516
 |||||
 QY 2653 GCTGTGCGGGCGGGGAGGGGTCGTCGGGGCGGCTACAGTCTCTCGGAAGGTCG 2712
 |||||
 Db 515 GCAGCTGCAGCTGC-----AGCTGCTCTCGTGGTACAGTCCCGCTAGCGGGTCTCA 462
 |||||
 QY 2713 AACATCTGGACGAAGGCGAGCTCTGTCAAGCGCATCTTGGCCCTTGGAGAGT 2772
 |||||
 Db 461 AACATCTGGACGAAGGCGAGCTCTGTCAATGGCATCTTGGCCCTTGGAGAGT 402
 |||||
 QY 2773 GACCTGGTGGGGCGGGGACTTTGGCAGCTCCGAATGGTGTATGGAGACGCGGTC 2832
 |||||
 Db 401 GACCTGGTGGGAGCTCGGGGACATTTGGCAGCTCAGAGTGGTATGGAGACTGTG 342
 |||||
 QY 2833 TAGGTGGGGGTGGGATGCTCCCTTTCTGTGCGCAGGSGTGGGAGAAAGAAATC- 2891
 |||||
 Db 341 TAGCAGGGAAGGATGACCGTCCCTCCCGCCCTTCTCAAAGTGGGAGAAAGGAGTCA 282
 |||||
 QY 2892 -----TACTGCGAAGTGTGTTGGAGTTCATGGTGTATGTATCATCCAGGACAGT 2945
 |||||
 Db 281 GAGTCAGCACTGGCAAGTATTTGTGGAGTTCCATGGTGTATGTATCATCCAGGACAGT 222
 |||||
 QY 2946 TTCGTCTCCCTGCAATGCCCTGTGTC-----CCCGCTACCCCGCAACCCACATC 2999
 |||||
 Db 221 CTTGTCTCCCTGCAACAGCTTATGTCCTTCTGCGCCAGCTACACCCCTTCCCTATGTC 162
 |||||
 QY 3000 ACCCTCCCGA-----CCACCGCGCGGGGTGTCTCAGGGAAATGGACTCGCTC 3048
 |||||
 Db 161 ACCCTCCCGAGGCTGTCACCCACCCAGCGGGTGTCTCAGGAAATG-GGACTCGCTC 103
 |||||
 QY 3049 AATGCCGACTGAGCCCTGAGTGTGAAAGGCGAGACTCGCCCTTTCTAATCACAAA 3108
 |||||
 Db 102 ATGTGTCAAGACTGAGCCCTGAGTGTGAAAGGCTTAACTGCTTTCTAATCACAAA 43
 |||||
 QY 3109 TGTAGCTCAAAGCAAGCGGCTTTGGA 3135
 |||||
 Db 42 TGTAG-CGACAAAGCGGCTTCGAA 17
 |||||

RESULT 14
 AW161909/c 397 bp mRNA linear EST 09-NOV-1999
 LOCUS au7le01.x1 Schneider fetal brain 00004 Homo sapiens cDNA clone
 DEFINITION IMAGE:2781720 3', mRNA sequence.
 ACCESSION AW161909
 VERSION AW161909.1 GI:6300942
 KEYWORDS EST
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 397)
 AUTHORS Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S.,
 Kizman, D., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin,
 J., Moore, B., Schellenberg, K., Steptoe, M., Tan, F., Theising, B.,
 White, Y., Wyllie, T., Waterston, R., Waterston, R. and Wilson, R.
 WASHU-NCI human EST Project
 Unpublished (1997)
 Other ESTs: au7le01.y1
 CONTACT: Wilson RK
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu
 This clone is available royalty-free through LILNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.

Seq primer: -40UP from Gibco
 High quality sequence stop: 390.
 Location/Qualifiers
 1..397
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:2781720"
 /clone_lib="Schneider fetal brain 00004"
 /sex="male"
 /tissue type="frontal lobe"
 /dev stage="5 months post-conception"
 /lab_host="DH10B"
 /note="Organ: brain; Vector: pBluescript SK (Stratagene);
 Site 1: SstI; Site 2: XhoI; Double-stranded cDNA was
 prepared from human fetal brain tissue. 5' and 3'
 adaptors were used in cloning as follows: 5' adaptor
 sequence:
 5'-GAGACAGAGAGAGCTCAAGGATCTTAATTAATAATTAATCCCGCCCGCC-3'
 and 3' adaptor sequence:
 5'-GAGACAGAGACTCGAGTCTTTTCTTTT-3'. The library was
 size-selected for >0.5 kb inserts and has an average
 insert size estimated at 1.2 kb. This library was
 constructed using the CAP-trapper method for full-length
 enrichment and has not undergone amplification. Library
 was constructed by Dr. Claudio Schneider (UNCIB-Area
 Science Park, Trieste, Italy)."

BASE COUNT 84 a 113 c 9 others
 ORIGIN
 Query Match 10.0%; Score 313; DB 10; Length 397;
 Best Local Similarity 90.6%; Pred. No. 2e-56; Indels 0; Gaps 0;
 Matches 328; Conservative 0; Mismatches 34;

QY 2781 GGGGGCCCGGGGACTTTTGGCAGCTCCGAATGGGTGATGGAGACACCGTCTAGGTGG 2840
 |||||
 Db 395 GGGGGCCCGGGGACTTTTGGCAGCTCCGAATGGGTGATGGAGACACCGTCTAGGTGG 336
 |||||
 QY 2841 GGTGGCATGCTCCCTTTCTGTGCGCAGGTTGGGAGAAAGGAAAGAAATCTACATGGCA 2900
 |||||
 Db 335 GGTGGCAAGCTCCCTTTCTGTGCGCAGGTTGGGAGAAAGGAAAGAAATCTACATGGCA 276
 |||||
 QY 2901 AGTGTGTTGGAGTTCATGCTGATGTTTACATCCAGGACAGTTCCTCTCCCTGTCA 2960
 |||||
 Db 275 AGGTTTGTGGAGTTCATGCTGATGTTTACCAACAGGACGGTTTTGTCTCCCGNNA 216
 |||||
 QY 2961 ATGGCTCGTGTCCCGCCCTACCCCGCAACACCCACATCACCTCCCGCCCGCCGG 3020
 |||||
 Db 215 AGGGCTGTGTCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCGG 156
 |||||
 QY 3021 GGTGTGCTCAGGAAATGTTGACTGCTCAAAATGCGCGACTGAGCCCTGAGTGTGTTGAAA 3080
 |||||
 Db 155 GGTGTGCTCAGGAAATGTTGACTGCTCAAAATGCGCGACTGAGCCCTGAGTGTGTTGAAA 96
 |||||
 QY 3081 GCGGACTCCCGCTTCTAATCACAAATGTAGCTACAAAGCGGCTTTGGATTGCT 3140
 |||||
 Db 95 GGGGAGACTCCCGCTTTTAAATCAAAAATGTAGCTTACAAAGCAAGGGGCTTTGGATTGTT 36
 |||||
 QY 3141 TA 3142
 |||||
 Db 35 TA 34
 |||||

RESULT 15
 BM046297
 LOCUS 603626095FI NIH_MGC_40 Homo sapiens cDNA clone IMAGE:5452630 5',
 DEFINITION mRNA sequence.
 ACCESSION BM046297
 VERSION BM046297.1 GI:16775564
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 1 (bases 1 to 708)
 REFERENCE NIH-MGC <http://mgs.nci.nih.gov/>
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Published: Robert Strausberg, Ph.D.
 Email: cgabbs-r@mail.nih.gov
 Tissue Procurement: DCTD/DTP
 CDNA Library Preparation: Ling Hong/Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: L1CM1944 row: n column: 23
 High quality sequence stop: 704.

FEATURES

Location/Qualifiers
 1..708
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5452630"
 /clone_lib="NIH_MGC_40"
 /tissue_type="carcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /notes="Organ: prostate; Vector: pOTB7; Site 1: XhoI;
 Site 2: EcoRI; cDNA made by oligo-dt priming.
 Directionally cloned into EcoRI/XhoI sites using the
 following 5' adaptor: GGCACGAG(G). Library constructed by
 Ling Hong in the laboratory of Gerald M. Rubin (University
 of California, Berkeley) using ZAP-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies).
 Note: this is a NIH MGC Library."

BASE COUNT 101 a 272 c 209 g 126 t
 ORIGIN
 Query Match 9.9%; Score 310.6; DB 13; Length 708;
 Best Local Similarity 65.8%; Pred. NO. 7e-56;
 Matches 473; Conservative 0; Mismatches 234; Indels 12; Gaps 1;

QY 651 CTTTCATCCACATCAGCGCCGAGGACTTCCACATGACGGGGCTGGTGGACCTGAC 710
 Db |||||||
 QY 2 CTTTCATCCAGGCCCTGGGGCCCTGCATTCGCGAACATGACGGGACTGGTGGACCTGAC 61
 Db |||||||
 QY 711 CTTGTCCAGGAACACCATCAGCCATCCAGCCCTTTCTTTCTGACCTCGAGAGCCT 770
 Db |||||||
 QY 62 ACTGTCGCAATGCCATCACCGCATTTGGGCCCGCGCTTTGGGACTCGAGAGCCT 121
 Db |||||||
 QY 771 CGCTCCCTGCAFTTTGACAGCAATCGGCTGCCAAGCCTTGGGGAGGACACCCCTCCGGGG 830
 Db |||||||
 QY 122 GCGTTCCTCCACTTTGAGGCAACAGGCTGTGGAGCTGGGCACCGGGAGCCTCCGGGG 181
 Db |||||||
 QY 831 CTTGGTCAACCTGACACCTTATCGTGAACACACACACAGCTGGGGGATCGCAGATGA 890
 Db |||||||
 QY 182 CCCCCTCAATTCGACGACCTTATCTCAGCGGCAACACAGCTGGGGCCGATCGGCCGGG 241
 Db |||||||
 QY 891 GGCTTTTGGAGCTTCTGTCACATTTGGAGGATCTGGACCTCTCCTACAACAACCTCCA 950
 Db |||||||
 QY 242 AGCCTTCGAGACTTCTAGAGAGCCTGGAGGACTGGACTGTCTCTACAAACCTCCG 301
 Db |||||||
 QY 951 TGGCCCTGGTGGGATCCCGTGGGACGCAATGFTCAAACCTCCACAGCTGAGCCTGGACCA 1010
 Db |||||||
 QY 302 GCAGGTGCCCTGGGGCCGATCGGCCCATGCTGCTGCACACCCCTCAAACCTGGACCA 361
 Db |||||||
 QY 1011 CAACCTGCTGGATCACATCGCGAGGGACCTTTGCAGACCTTCAGAACTGGCCGGCCT 1070
 Db |||||||
 QY 362 TAACTTATGAGGACTGCCCCAGGCCCTTTGGCCAGCTCGGTCAGCTCTCCGCCCT 421
 Db |||||||
 QY 1071 GGATCTCACCTCCAATCGGCTGCAGAAAGTGGCCCTGATCCCATCTTTGGCCGCTCCA 1130
 Db |||||||
 QY 422 GGACCTCACCTCAACCGCCTGSCCAGCTGGCTCGGACCCGCTTTCTCTCGTGGGG 481
 Db |||||||
 QY 1131 GCGTTGGGCTTTGACAGCACACCCCTTTGCCACCCCTTGTCTTTAGTTTGGGGGTTAA 1190
 Db |||||||

Db 482 TGATGCAGAGG-----CCTCTCCCGCCCGCCCTGTGTGCTGAGCTTTAGCGGGAA 529
 QY 1191 CCCACTTCACTGCAATTTGTGAGCTTCTCTGCTGCGGAGGCTCCGAGCGGGAGGATGACT 1250
 Db |||||||
 QY 530 CCCCCTGCACTGCAAACTGTGAGCTGTGTGCTGCGCGGCTGGCGCCGCGGACGACT 589
 Db |||||||
 QY 1251 GGAAACCTGTGGCTCCCCCAGGGGCTCAAGGGTCCGTACTTCTGGCATGTGCGTGAGGA 1310
 Db |||||||
 QY 590 GGAAACGTTGGCCCTCCCGACCGGCTGCGCCGCGCTACTTCTGGGCAAGTGGCCGAGGG 649
 Db |||||||
 QY 1311 GGAGTTTGTGCGGAGCCGCTTCATCACCCAGCACACACAAAGTTGCTGGTTCTGG 1369
 Db |||||||
 QY 650 CGAGTTCTCTGTGAGCCTGGCCCTCATTTGCCCGACACACCGAGCCCTCTGGGTGCTTG 708
 Db |||||||

Search completed: June 8, 2003, 19:02:56
 Job time : 3928 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 8, 2003, 19:03:03 ; Search time 63 Seconds
(without alignments)
1668.805 Million cell updates/sec

Title: US-09-831-846-2
Perfect score: 4094
Sequence: 1 METLLGGLLAFGMFAVVDAA.....DLVGARFTGSGSEWVMEVTV 789

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : A_Geneseq_101002.*
1: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT.*
23: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Table with columns: Result NO., Score, Match, Length, DB ID, Description. Contains 23 rows of search results.

Table with columns: ID, Score, Match, Length, DB ID, Description. Contains 45 rows of search results.

ALIGNMENTS

RESULT 1
AAB09968
ID AAB09968 standard; Protein; 789 AA.
AC AAB09968;
DT 19-OCT-2000 (first entry)
XX Human brain-specific transmembrane glycoprotein.
DE Transmembrane glycoprotein; human; brain; cellular signal transducer;
KW neuroactive; neuroprotective; cerabroprotective; drug development;
KW treatment; nervous disease; diagnostic.
XX Homo sapiens.
OS WO2000031256-A1.
XX 02-JUN-2000.
XX 18-NOV-1999; 99WO-JP06449.
XX 20-NOV-1999; 98JP-0331727.
XX (CHUG-) CHUGAI RES INST MOLECULAR MEDICINE INC.
XX (KAZU-) KAZUSA DNA RES INST.
XX Funahashi S, Miyata S, Nomura N, Nagase T, Ohata O;
XX WPI: 2000-411954/35.
XX N-PSDB; AAA40083.
XX Gene encoding a brain-specific transmembrane glycoprotein that has a

PT typical PDZ protein binding motif and functions as a cellular signal transducer, useful in developing drugs for treating nervous diseases -

XX Claim 1a; Page 52-57; 63pp; Japanese.

CC This invention describes a novel gene encoding a human brain-specific transmembrane glycoprotein that has a typical PDZ protein binding motif and functions as a cellular signal transducer. The product of the invention has neuroactive, neuroprotective and cerebroprotective activity. The gene and encoded protein are useful in developing drugs for treating nervous diseases, and also for studying functions of the nervous system or onset mechanism of nerve-related diseases. They may also be used for the production of diagnostic reagents such as oligonucleotide probes and antibodies for detecting proteins comprising PDZ domains and the nucleic acids that encode them. They may also be used to produce and identify modulators of the proteins expression and activity such as antisense sequences and antibodies. This sequence represents the human brain-specific transmembrane glycoprotein described in the method of the invention.

XX Sequence 789 AA;

Query Match 100.0%; Score 4094; DB 21; Length 789;

Best Local Similarity 100.0%; Pred No. 1.8e-303;

Matches 789; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 METLLGGLLAFGMAFAVVDACPKYCVQNLSESLGTLCPKGLLFFVPPIDRRIVLRLG 60

DB 1 METLLGGLLAFGMAFAVVDACPKYCVQNLSESLGTLCPKGLLFFVPPIDRRIVLRLG 60

QY 61 GNFIHHSRODFANMTGLVLTLSRNTISHIQPPSFLDLSLHLSNRLPSLGEDTL 120

DB 61 GNFIHHSRODFANMTGLVLTLSRNTISHIQPPSFLDLSLHLSNRLPSLGEDTL 120

QY 121 RGLVNLQHLIWNVNLGGIADEAFEDFLLTLEDLSDYNNLHGLPWSVRRMVLHQLSL 180

DB 121 RGLVNLQHLIWNVNLGGIADEAFEDFLLTLEDLSDYNNLHGLPWSVRRMVLHQLSL 180

QY 181 DHNLLDHIAGTFADLQKARLDLTSNRLOKLPDDPIFARSOASALTATFAPPLSFSFG 240

DB 181 DHNLLDHIAGTFADLQKARLDLTSNRLOKLPDDPIFARSOASALTATFAPPLSFSFG 240

QY 241 GNPLHNCCELLWLRRLRDEDDLETGSPGGLKGRYFHWVREEFVCEPPLITQHTKLLV 300

DB 241 GNPLHNCCELLWLRRLRDEDDLETGSPGGLKGRYFHWVREEFVCEPPLITQHTKLLV 300

QY 301 LEGQAATLKCAIGDPPSLIHWAPDRLVGNSSRTAVYDNGTLDIFITTSQDSGAFTCI 360

DB 301 LEGQAATLKCAIGDPPSLIHWAPDRLVGNSSRTAVYDNGTLDIFITTSQDSGAFTCI 360

QY 361 AANAAGEATAMVEVSIVQLPHLSNSTSTAPPKSLSDITGSSKTSRGGGGGGGEPKPS 420

DB 361 AANAAGEATAMVEVSIVQLPHLSNSTSTAPPKSLSDITGSSKTSRGGGGGGGEPKPS 420

QY 421 PPERAVLVEVTTTSALVKWSKSAKPRKMYQVYQVYQVYQVYQVYQVYQVYQVYQV 480

DB 421 PPERAVLVEVTTTSALVKWSKSAKPRKMYQVYQVYQVYQVYQVYQVYQVYQV 480

QY 481 LVSGTGYDLCVLAAMDDTATLTATNIVGCAQFFTKADYPOCOSMHSQILGGTMLIVIG 540

DB 481 LVSGTGYDLCVLAAMDDTATLTATNIVGCAQFFTKADYPOCOSMHSQILGGTMLIVIG 540

QY 541 IIVATLLFVIVLWRYKVCNHEAPSKMAAAVSNVYQVYQVYQVYQVYQVYQVYQVYQV 600

DB 541 IIVATLLFVIVLWRYKVCNHEAPSKMAAAVSNVYQVYQVYQVYQVYQVYQVYQV 600

QY 601 KVVVRNELLDFTLASLARSDDSSSSSLGSGEAAAGLGRAPWRIPPSAPRKPSSLRLMGAF 660

DB 601 KVVVRNELLDFTLASLARSDDSSSSSLGSGEAAAGLGRAPWRIPPSAPRKPSSLRLMGAF 660

QY 661 ASLDLKSQRKELLDSRTAGAGTASARGHSDREPLLGPAAARARSLLPFLGKAKR 720

DB 661 ASLDLKSQRKELLDSRTAGAGTASARGHSDREPLLGPAAARARSLLPFLGKAKR 720

Qy 721 SHSFDMGDFRAAAGGVPVGGYSPPRKVSNIWTKRSLSVNGMLLPFEESDLVGARGTFGS 780

Db 721 SHSFDMGDFRAAAGGVPVGGYSPPRKVSNIWTKRSLSVNGMLLPFEESDLVGARGTFGS 780

Qy 781 SEWMESTV|789

Db 781 SEWMESTV|789

RESULT 2

AAU28092

ID AAU28092 standard; Protein; 789 AA.

XX AAU28092;

AC AAU28092;

XX 18-DEC-2001 (first entry)

XX Novel human secretory protein, Seq ID No 261.

XX Human; secreted protein; arthritis; Crohn's disease; sepsis; shock; ischaemia-reperfusion injury; haematopoiesis; cancer; neuropathy; transgenic animal; Alzheimer's disease; Parkinson's disease; burn; amyotrophic lateral sclerosis; platelet disorder; thrombocytopenia; ulcer; osteoporosis; bone degenerative disorder; periodontal disease; gut protection; lung; liver fibrosis; immune deficiency; infection; severe combined immunodeficiency; SCID; autoimmune disorder; allergy; multiple sclerosis; rheumatoid arthritis; diabetes mellitus; asthma; fertility; analgesic; pain; antigen.

XX Homo sapiens.

OS WO200166689-A2.

XX 13-SEP-2001.

XX 05-MAR-2001; 2001WO-0504942.

XX 07-MAR-2000; 2000US-0519705.

XX 19-MAY-2000; 2000US-0574454.

XX 17-JUN-2000; 2000US-0596193.

XX 14-JUL-2000; 2000US-0616847.

XX 19-SEP-2000; 2000US-0665363.

XX 20-OCT-2000; 2000US-0693267.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Asundi V, Xu C, Wehrman T, Ren F, Ma Y, Zhou P;

XX Zhao QA, Yang Y, Drmanac RT, Zhang J, Chen R, Xue AJ, Wang J;

XX WPI; 2001-589934/66.

XX N-PSDB; AAS44992.

XX Novel polypeptides and nucleic acids obtained from cDNA libraries prepared from various human tissues, for diagnosis and treatment of cancer, neurological, inflammatory, and autoimmune disorders.

XX Example 4; SEQ ID No 261; 107pp; English.

XX The invention relates to novel isolated human secreted polypeptides (I) and polynucleotides (II). (I) and (II) are useful for treating inflammatory conditions such as arthritis, nephritis, Crohn's disease, ischaemia-reperfusion injury, shock, sepsis, immune responses, and is involved in increasing haematopoiesis, stem cell survival, bone growth and remodeling. (I), (II) and modulators of (II) are useful for Crohn's disease or treatment of one or more cancers. (II) is also useful for creating transgenic animals useful for studying the in vivo activities of the polypeptide as well as for studying modulators of neuronal cells and regeneration of nerve (I) induces the proliferation of neural cells and regeneration of nerve CC (I) induces the proliferation of neural cells and regeneration of nerve CC peripheral nervous system diseases and neuropathies, such as Alzheimer's, CC Parkinson's disease, Huntington's disease, and amyotrophic lateral sclerosis. In addition, (I) is involved in chemotactic or chemokinetic

CC activity, regulation of haematopoiesis and is useful for treating myeloid
 CC or lymphoid cell disorders, platelet disorders such as thrombocytopenia
 CC and for regeneration of bone, cartilage, tendon, ligament and/or nerve
 CC tissue growth, and in tissue repair, healing of burns, incisions,
 CC ulcers, for treating osteoporosis, osteoarthritis, bone degenerative
 CC disorders or periodontal disease. Furthermore (I) is also useful for
 CC sut protection or regeneration and treatment of lung or liver fibrosis,
 CC reperfusion injury in various tissues, various immune deficiencies and
 CC disorders including severe combined immunodeficiency (SCID), bacterial or
 CC fungal infections, autoimmune disorders e.g. multiple sclerosis,
 CC rheumatoid arthritis, diabetes mellitus, myasthenia gravis, allergic
 CC reactions and conditions, such as asthma or other respiratory problems.
 CC In addition, (I) affects biorhythms or circadian cycles of rhythms,
 CC fertility, metabolism, catabolism, anabolism, storage or elimination of
 CC dietary fat, lipid, protein, carbohydrate, vitamins, minerals, provides
 CC analgesic effects or other pain reducing effects, immunoglobulin like
 CC activity and can act as an antigen in a vaccine composition to raise an
 CC immune response. AAU28020-AAU28395 represent novel human secreted protein
 CC amino acid sequences of the invention.
 XX

Sequence 789 AA;

Query Match 99.8%; Score 4087; DB 22; Length 789;
 Best Local Similarity 99.9%; Pred. No. 66-303; 1; Indels 0; Gaps 0;
 Matches 788; Conservative 0; Mismatches 1;

Qy	1	METLLGGLAFGMFAVVDACPKYCVQNLSESLGTLCPKGLLFPVDPIDRRTVELRLG	60
Db	1	METLLGGLAFGMFAVVDACPKYCVQNLSESLGTLCPKGLLFPVDPIDRRTVELRLG	60
Qy	61	GNFIIHSRQDPANMTGLVDLTLRNITSHIQPFSLDESRLHLSNRLPSLGEDTL	120
Db	61	GNFIIHSRQDPANMTGLVDLTLRNITSHIQPFSLDESRLHLSNRLPSLGEDTL	120
Qy	121	RGVNLQHLIVNNQLGGTADAFEDFLTLEDLDSYNNLHGLPWSVRRMVLHQLSL	180
Db	121	RGVNLQHLIVNNQLGGTADAFEDFLTLEDLDSYNNLHGLPWSVRRMVLHQLSL	180
Qy	181	DHNLDHIAEGTFADLQKLARLDLTSNRLOKLPDPPIFARSQASALTATFPAPLFSFG	240
Db	181	DHNLDHIAEGTFADLQKLARLDLTSNRLOKLPDPPIFARSQASALTATFPAPLFSFG	240
Qy	241	GNPLHCNCELLWRLERDDDLTCGSPGGLKGRYFHWVREEFVCEPPLITQHTKLIV	300
Db	241	GNPLHCNCELLWRLERDDDLTCGSPGGLKGRYFHWVREEFVCEPPLITQHTKLIV	300
Qy	301	LEGQAATLKCKAIGDPSPLIHWVAPDRLVGNSSRTAVYDNGTLDIFITTSODSGAFTCI	360
Db	301	LEGQAATLKCKAIGDPSPLIHWVAPDRLVGNSSRTAVYDNGTLDIFITTSODSGAFTCI	360
Qy	361	AANAAGEATAMVEVSIVQLPHLSNSTRTPPKSRSLDITGSSKTSRGGSGGGEPPPKS	420
Db	361	AANAAGEATAMVEVSIVQLPHLSNSTRTPPKSRSLDITGSSKTSRGGSGGGEPPPKS	420
Qy	421	PPERAVLVSEVTTTALSALVWVSKSAPRVKMYQLQVNCSDDEVLIYRMIIPASNKA	480
Db	421	PPERAVLVSEVTTTALSALVWVSKSAPRVKMYQLQVNCSDDEVLIYRMIIPASNKA	480
Qy	481	LVSGTGYDLVCLAMWDDTATTLTATNIVGCAQFFTKADYPOCSMHSQILGGTMIIVIGG	540
Db	481	LVSGTGYDLVCLAMWDDTATTLTATNIVGCAQFFTKADYPOCSMHSQILGGTMIIVIGG	540
Qy	541	IIVATLLVFIIVLWVRYKVCNHEAPSKMAAAVSNVYQTNQAQPPPPSAPAGAPQGGP	600
Db	541	IIVATLLVFIIVLWVRYKVCNHEAPSKMAAAVSNVYQTNQAQPPPPSAPAGAPQGGP	600
Qy	601	KVVVRNELLDFATSLARASDSSSSSLGSGEAGLGRAPWRIIPSPAPRPKSLDLRMGAF	660
Db	601	KVVVRNELLDFATSLARASDSSSSSLGSGEAGLGRAPWRIIPSPAPRPKSLDLRMGAF	660
Qy	661	ASLDLKSQRKEELLSRTPAGRGAGTSARGHHSREPLLGPAAARALLPLPLSGKAKR	720
Db	661	ASLDLKSQRKEELLSRTPAGRGAGTSARGHHSREPLLGPAAARALLPLPLSGKAKR	720

Qy	721	SHSFDMDGFAAAAAAGVVVGGYSPRKYVNIWTKRSLVNGMLLPFEESDLVARGTFCG	780
Db	721	SHSFDMDGFAAAAAAGVVVGGYSPRKYVNIWTKRSLVNGMLLPFEESDLVARGTFCG	780
Qy	781	SEWVMESTV 789	
Db	781	SEWVMESTV 789	
RESULT 3			
AA039059	standard; Protein; 789 AA.		
XX	AA039059;		
XX	22-OCT-2001 (first entry)		
XX	Human polypeptide SEQ ID NO 2204.		
XX	Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;		
KW	peripheral nervous system; neuropathy; central nervous system; CNS;		
KW	Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;		
KW	amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;		
KW	chemokinetic; thrombolytic; drug screening; arthritis; inflammation;		
KW	leukaemia.		
XX	Homo sapiens.		
XX	WO200153312-A1.		
XX	26-JUL-2001.		
XX	26-DEC-2000; 2000WO-US34263.		
XX	21-JAN-2000; 2000US-0488725.		
PR	25-APR-2000; 2000US-0552317.		
PR	09-JUL-2000; 2000US-0598042.		
PR	19-JUL-2000; 2000US-0620312.		
PR	03-AUG-2000; 2000US-0653450.		
PR	14-SEP-2000; 2000US-0662191.		
PR	19-OCT-2000; 2000US-0693036.		
PR	29-NOV-2000; 2000US-0727344.		
XX	(HYSE-) HYSEQ INC.		
XX	Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;		
PI	Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;		
PI	Zhao QA, Zhou P, Goodrich R, Drmanac RT;		
XX	WPI; 2001-442253/47.		
DR	N-PSDB; AAI58215.		
XX	Novel nucleic acids and polypeptides, useful for treating disorders		
PT	such as central nervous system injuries -		
XX	Example 4; SEQ ID NO 2204; 10078pp; English.		
XX	The invention relates to human nucleic acids (AAI57798-AAI61369) and		
CC	the encoded polypeptides (AAM38642-AAM42213) with nootropic,		
CC	immunosuppressant and cytostatic activity. The polynucleotides are useful		
CC	in gene therapy. A composition containing a polypeptide or polynucleotide		
CC	of the invention may be used to treat diseases of the peripheral nervous		
CC	system, such as peripheral nervous injuries, peripheral neuropathy and		
CC	localised neuropathies and central nervous system diseases, such as		
CC	Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic		
CC	lateral sclerosis, and Shy-Drager Syndrome. Other uses include the		
CC	utilisation of the activities such as: Immune system suppression,		
CC	Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic		
CC	and thrombolytic activity, cancer diagnosis and therapy, drug screening,		
CC	assays for receptor activity, arthritis and inflammation, leukaemias and		
CC	C.N.S disorders.		
CC	Note: The sequence data for this patent did not form part of the printed		

CC specification.
 XX Sequence 789 AA;
 SQ Query Match 99.8%; Score 4087; DB 22; Length 789;
 Best Local Similarity 99.9%; Pred No. 6e-303; Indels 0; Gaps 0;
 Matches 788; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

DE Human hh00149 protein SEQ ID NO:4.
 KW Human; ubiquitin-like protein; 149Y2H#151; hh00149; brain;
 KW 2-hybrid screening; neuroprotective; signal transducer;
 KW nervous system disease; diagnosis.
 XX Homo sapiens.
 XX W0200031255-A1.
 XX 02-JUN-2000.
 XX 18-NOV-1999; 99WO-JP06448.
 XX 20-NOV-1998; 98JP-0331701.
 XX (CHUG-) CHUGAI RES INST MOLECULAR MEDICINE INC.
 XX Funahashi S, Miyata S;
 XX WPI; 2000-400066/34.
 XX N-PSDB; AAA60605.
 XX Gene encoding ubiquitin-like protein which interacts with protein
 hh00149, useful in the diagnosis and treatment of diseases associated
 with the nervous system -
 XX Example 2; Page 74-79; 88pp; Japanese.
 XX The present invention describes a ubiquitin-like protein, designated
 149Y2H#151, which interacts with protein hh00149 expressed specifically
 in the brain. The 149Y2H#151 protein has neuroprotective activity, and
 is a signal transducer. The 149Y2H#151 gene and encoded protein are
 useful in the diagnosis and treatment of diseases associated with the
 nervous system. The protein can interact with protein hh00149 expressed
 specifically in the brain. The present sequence represents the human
 hh00149 protein.

5 LGGLAFCGMAFAVVDACPKYCVQNLSLGLTCLPCKGLLFPDPIDRRRTVELRGGNFI 64
 1 LGGLAFCGMAFAVVDACPKYCVQNLSLGLTCLPCKGLLFPDPIDRRRTVELRGGNFI 60
 65 IHSRQDFANMTGLVDLTLRSNTISHIQFSPFLDLESRLSLHLDNSRLPSLGEDTLRGLV 124
 61 IHSRQDFANMTGLVDLTLRSNTISHIQFSPFLDLESRLSLHLDNSRLPSLGEDTLRGLV 120
 125 NLOHLIVNNQLGGIADEAFEDLTLLEDLDSYNNHLHGLPWDSVRRMVNHLQSLDHNH 184
 121 NLOHLIVNNQLGGIADEAFEDLTLLEDLDSYNNHLHGLPWDSVRRMVNHLQSLDHNH 180
 185 LDHIAEGTADLQKARLDLTSNRLOKLPDDIFARSOASALTATPFAPPLSFSGGNPL 244
 181 LDHIAEGTADLQKARLDLTSNRLOKLPDDIFARSOASALTATPFAPPLSFSGGNPL 240
 245 HCNCELLWLRRLERDDDLTCCSGPGKGRYFHWVREBEFVCEPPLITQHTKLLVLEGO 304
 241 HCNCELLWLRRLERDDDLTCCSGPGKGRYFHWVREBEFVCEPPLITQHTKLLVLEGO 300
 305 AATLKCAIKGDFSPLIHWVAPDDRLVGNSSRTAVYDNGTLDIFITTSQDSGAFCTAANA 364
 301 AATLKCAIKGDFSPLIHWVAPDDRLVGNSSRTAVYDNGTLDIFITTSQDSGAFCTAANA 360
 365 AGEATAMVEVSIVQLPHLSNSTRTAPPKSRSLDITGSSKTSRGGGCGGPPKPPPP 424
 361 AGEATAMVEVSIVQLPHLSNSTRTAPPKSRSLDITGSSKTSRGGGCGGPPKPPPP 420
 425 AVLVSEVTTTTSALVKWSVSKSAPRVKMYQLQYNCSDDEVLIYRMI PASNKAFVYNNLVSG 484

Query Match 99.5%; Score 4075; DB 21; Length 785;
 Best Local Similarity 100.0%; Pred. No. 4.9e-302;
 Matches 785; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 LGGLAFCGMAFAVVDACPKYCVQNLSLGLTCLPCKGLLFPDPIDRRRTVELRGGNFI 64
 Db 1 LGGLAFCGMAFAVVDACPKYCVQNLSLGLTCLPCKGLLFPDPIDRRRTVELRGGNFI 60
 QY 65 IHSRQDFANMTGLVDLTLRSNTISHIQFSPFLDLESRLSLHLDNSRLPSLGEDTLRGLV 124
 Db 61 IHSRQDFANMTGLVDLTLRSNTISHIQFSPFLDLESRLSLHLDNSRLPSLGEDTLRGLV 120
 QY 125 NLOHLIVNNQLGGIADEAFEDLTLLEDLDSYNNHLHGLPWDSVRRMVNHLQSLDHNH 184
 Db 121 NLOHLIVNNQLGGIADEAFEDLTLLEDLDSYNNHLHGLPWDSVRRMVNHLQSLDHNH 180
 QY 185 LDHIAEGTADLQKARLDLTSNRLOKLPDDIFARSOASALTATPFAPPLSFSGGNPL 244
 Db 181 LDHIAEGTADLQKARLDLTSNRLOKLPDDIFARSOASALTATPFAPPLSFSGGNPL 240
 QY 245 HCNCELLWLRRLERDDDLTCCSGPGKGRYFHWVREBEFVCEPPLITQHTKLLVLEGO 304
 Db 241 HCNCELLWLRRLERDDDLTCCSGPGKGRYFHWVREBEFVCEPPLITQHTKLLVLEGO 300
 QY 305 AATLKCAIKGDFSPLIHWVAPDDRLVGNSSRTAVYDNGTLDIFITTSQDSGAFCTAANA 364
 Db 301 AATLKCAIKGDFSPLIHWVAPDDRLVGNSSRTAVYDNGTLDIFITTSQDSGAFCTAANA 360
 QY 365 AGEATAMVEVSIVQLPHLSNSTRTAPPKSRSLDITGSSKTSRGGGCGGPPKPPPP 424
 Db 361 AGEATAMVEVSIVQLPHLSNSTRTAPPKSRSLDITGSSKTSRGGGCGGPPKPPPP 420
 QY 425 AVLVSEVTTTTSALVKWSVSKSAPRVKMYQLQYNCSDDEVLIYRMI PASNKAFVYNNLVSG 484

CC specification.
 XX Sequence 789 AA;
 SQ Query Match 99.8%; Score 4087; DB 22; Length 789;
 Best Local Similarity 99.9%; Pred No. 6e-303; Indels 0; Gaps 0;
 Matches 788; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 METLLGGLAFCGMAFAVVDACPKYCVQNLSLGLTCLPCKGLLFPDPIDRRRTVELRGLG 60
 Db 1 METLLGGLAFCGMAFAVVDACPKYCVQNLSLGLTCLPCKGLLFPDPIDRRRTVELRGLG 60
 QY 61 GNPIIHSRQDFANMTGLVDLTLRSNTISHIQFSPFLDLESRLSLHLDNSRLPSLGEDTL 120
 Db 61 GNPIIHSRQDFANMTGLVDLTLRSNTISHIQFSPFLDLESRLSLHLDNSRLPSLGEDTL 120
 QY 121 RGLVNLQHLIVNNQLGGIADEAFEDLTLLEDLDSYNNHLHGLPWDSVRRMVNHLQSL 180
 Db 121 RGLVNLQHLIVNNQLGGIADEAFEDLTLLEDLDSYNNHLHGLPWDSVRRMVNHLQSL 180
 QY 181 DHNLLDHI AEGTADLQKARLDLTSNRLOKLPDDIFARSOASALTATPFAPPLSFSG 240
 Db 181 DHNLLDHI AEGTADLQKARLDLTSNRLOKLPDDIFARSOASALTATPFAPPLSFSG 240
 QY 241 GNPLHCNCELLWLRRLERDDDLTCCSGPGKGRYFHWVREBEFVCEPPLITQHTKLLV 300
 Db 241 GNPLHCNCELLWLRRLERDDDLTCCSGPGKGRYFHWVREBEFVCEPPLITQHTKLLV 300
 QY 301 LEGOATLKCAIKGDFSPLIHWVAPDDRLVGNSSRTAVYDNGTLDIFITTSQDSGAFCTI 360
 Db 301 LEGOATLKCAIKGDFSPLIHWVAPDDRLVGNSSRTAVYDNGTLDIFITTSQDSGAFCTI 360
 QY 361 AANAAGATAMVEVSIVQLPHLSNSTRTAPPKSRSLDITGSSKTSRGGGCGGPPK 420
 Db 361 AANAAGATAMVEVSIVQLPHLSNSTRTAPPKSRSLDITGSSKTSRGGGCGGPPK 420
 QY 421 PPERAVLVSEVTTTTSALVKWSVSKSAPRVKMYQLQYNCSDDEVLIYRMI PASNKAFV 480
 Db 421 PPERAVLVSEVTTTTSALVKWSVSKSAPRVKMYQLQYNCSDDEVLIYRMI PASNKAFV 480
 QY 481 LVSCTGYDLCVLA MWDDTATLTATNIVGCAQFFTKADYPCQCSMHSQILGGTMILVIG 540
 Db 481 LVSCTGYDLCVLA MWDDTATLTATNIVGCAQFFTKADYPCQCSMHSQILGGTMILVIG 540
 QY 541 IIVATLLVFIIVLVRVKVNCHEAPSKMAAAVSNVYOTNGAOPPPSSAPAGPOGPP 600
 Db 541 IIVATLLVFIIVLVRVKVNCHEAPSKMAAAVSNVYOTNGAOPPPSSAPAGPOGPP 600
 QY 601 KVVVRNELLDFASLARA SSSSSSSLSGSGEAAGLGRAPWRI PPSAPRPKPSLDRMLGAF 660
 Db 601 KVVVRNELLDFASLARA SSSSSSSLSGSGEAAGLGRAPWRI PPSAPRPKPSLDRMLGAF 660
 QY 661 ASLDLKSQRKEBELLSRTPAGRGAGTSARGHHSREPLLPAPPAARASLPLPLEGKAKR 720
 Db 661 ASLDLKSQRKEBELLSRTPAGRGAGTSARGHHSREPLLPAPPAARASLPLPLEGKAKR 720
 QY 721 SHSFDMDGFAAAAGVVPGGYPPKVSNTWTKRSLVNGMLLPPFESDLVARGTFGS 780
 Db 721 SHSFDMDGFAAAAGVVPGGYPPKVSNTWTKRSLVNGMLLPPFESDLVARGTFGS 780
 QY 781 SEWVMESTV 789
 Db 781 SEWVMESTV 789

RESULT 4
 AAB12448
 ID AAB12448 standard; Protein; 785 AA.
 XX
 AC AAB12448;
 XX
 DT 19-OCT-2000 (first entry)
 XX

Db 421 AVLVSEVTTTALVKWVSKSAPRVKMYQLQVNCSDDEVLIRMIIPASNKAFVYNNLVSG 480
 QY 485 TGYDLCLVAMWDDTATLTLATNIVGCAQFFTRADYPOCOSMHSOILGGTMIIVIGIIVA 544
 Db 481 TGYDLCLVAMWDDTATLTLATNIVGCAQFFTRADYPOCOSMHSOILGGTMIIVIGIIVA 540
 QY 545 TLLVFIIVLMVRYKVCNHEAPSKMAAAVSNVYSQTNGAQQPPPPSSAPAGAPQGGPKVVV 604
 Db 541 TLLVFIIVLMVRYKVCNHEAPSKMAAAVSNVYSQTNGAQQPPPPSSAPAGAPQGGPKVVV 600
 QY 605 RNELLDFTASLARASDSSSSSLGSGEAAAGLGRAPWRIIPSPAPRPKPSIDRLMGAFASLD 664
 Db 601 RNELLDFTASLARASDSSSSSLGSGEAAAGLGRAPWRIIPSPAPRPKPSIDRLMGAFASLD 660
 QY 665 LKSQRKEELLDRTPAGRCAGTARSAGHSDREPLGPPAARARSLLPLLEGKAKRSHSF 724
 Db 661 LKSQRKEELLDRTPAGRCAGTARSAGHSDREPLGPPAARARSLLPLLEGKAKRSHSF 720
 QY 725 DMGDFAAAAAGGVPGYSPRRKVSNIWTKRSLSVNGMLLPFEESDLVGARGTFGSSVWV 784
 Db 721 DMGDFAAAAAGGVPGYSPRRKVSNIWTKRSLSVNGMLLPFEESDLVGARGTFGSSVWV 780
 QY 785 MESTV 789
 Db 781 MESTV 785

RESULT 5
 AAG67505
 ID AAG67505 standard; Protein; 766 AA.
 AC AAG67505;
 XX
 DT 26-NOV-2001 (first entry)
 DE Amino acid sequence of a human secreted polypeptide.
 XX
 KW Human; secreted polypeptide; nervous disease; muscular disease; tumour;
 KW gastrointestinal ulceration; spinal cord disease; trachea disease;
 KW thyroid gland disease; ovary disease; prostate disease; heart disease;
 KW renal gland disease; small intestine disease; thymus disease;
 KW lymph node disease; muscular system disease; colon disease;
 KW lipase deficiency; cystic fibrosis; pancreatitis; clot formation;
 KW myocardial infarction; angioplasty; liver disease; coagulation disorder;
 KW microbial disease; immune disorder; inflammation; transplant rejection;
 KW bone thickness; bone density; ferroxidase loss; apoptosis;
 KW vascular smooth cell proliferation; vaccine.
 XX
 OS Homo sapiens.
 XX
 PN WO2001166690-A2.
 XX
 PD 13-SEP-2001.
 XX
 PF 05-MAR-2001; 2001WO-US07143.
 XX
 PR 06-MAR-2000; 2000US-0187107.
 PR 13-MAR-2000; 2000US-0188916.
 PR 03-OCT-2000; 2000US-0236874.
 PR 03-OCT-2000; 2000US-0237846.
 XX
 PA (SMIK) SMITHKLINE BEECHAM CORP.
 PA (SMIK) SMITHKLINE BEECHAM PLC.
 XX
 PI Agarwal P, Murdoch PR, Rizvi SK, Smith RF, Xiang Z;
 XX
 DR WPI; 2001-570768/64.
 DR N-PSDB; AAH78197.
 XX
 PT Novel isolated secreted polypeptide useful for treating nervous and
 PT muscular diseases, gastrointestinal ulceration, coagulation and immune
 PT disorders, microbial diseases, inflammation and transplant rejection

XX Claim 1; Page 58-60; 102pp; English.
 PS
 CC The present sequence represents a human secreted polypeptide. The
 CC secreted polypeptides and polynucleotides are useful for treating
 CC nervous and muscular diseases, for inhibiting tumour formation and
 CC metastasis, for treating gastrointestinal ulceration, for preventing
 CC and treating diseases in spinal cord, thyroid gland, ovary, prostate,
 CC renal gland, small intestine, heart, trachea, thymus, lymph node,
 CC muscular system and colon, for treating lipase deficiency in cystic
 CC fibrosis and pancreatitis, for treating undesirable clot formation
 CC such as myocardial infarction, during angioplasty and all surgical
 CC procedures that require decreased blood clot formation, for treating
 CC liver diseases, coagulation disorders and microbial diseases, for
 CC treating immune disorders, for treating inflammation and transplant
 CC rejection, for enhancing bone thickness and increasing bone density,
 CC for reducing the loss of essential ferroxidases, for suppressing
 CC apoptosis, and for regulating vascular smooth cell proliferation. They
 CC may also be used as vaccines.
 XX
 SQ Sequence 766 AA;
 Query Match 45.1%; Score 1846; DB 22; Length 766;
 Best Local Similarity 48.5%; Pred. No. 5.3e-132;
 Matches 383; Conservative 128; Mismatches 225; Indels 54; Gaps 11;
 QY 1 METLLGGLLAFGMFAVVDACPKYVCQNLSESLGTLCPKGLLFPVPPIDRRITVELRLG 60
 Db 1 MEKVLFLYFLGIA-VKAIQICPKRCVQQLSPLNATLCAKKGLLFVPPNIDRTVELRLA 59
 QY 61 GNFIHISQDFANMTGLVDLTLRSNTISHIOPFFLDLESRLSLHLDNRNLPUGEDTL 120
 Db 60 DNFVTNIKRKDFANMTSLVDLTLRSNTISFIHPAFADLRNLRALHLSNRLTKITNDFM 119
 QY 121 RGVNLHLLVNNLGGTADAEAFEDFLTLLEDLDSYNNLHGLPWSVRRVWNLHLSL 180
 Db 120 SGLSNLHLLVNNLGGTADAEAFEDFLTLLEDLDSYNNLHGLPWSVRRVWNLHLSL 178
 QY 181 DHNLDPHIAEGTFADLQKLRDLTSLNRLQKLPDPPIFARSOASALATATFPAPLPSFSG 240
 Db 179 DHNMIDNIPKGFSLHKKWLDVTSNKLQKLPDPPLFORAQLVATSGIISPTFALSFG 238
 QY 241 GNPLHCNCELLWRLRLRDDLETCGSPGGLKGRVFWVREBEFVCEPPLIQTHTKLLV 300
 Db 239 GNPLHCNCELLWRLRLRDDLETCASPPLLTGRYFWSIPEBEFCEPPLIQRHTRHEMRV 298
 QY 301 LEGQATLKCKAIGPSPPLIHWVAPDDRLVGNSSRTAVYDNGTLDLFIITSDSGAFTCI 360
 Db 299 LEGQATLRCKARGDPEPAIHWISPEGKLISNATRSVYDNGTLDLFIITVXDGTGFTCI 358
 QY 361 AANAAGEATAMVEVSTVOLPHLSNSTRAPPKRSRLSDITGSSKTSRGGGGGGGPPPKS 420
 Db 359 ASNPAGEATQIVDLHIIKPLHLLNHNHIEPDPGSSDLSITKSGSNSTSSNG---DTK 415
 QY 421 PPERAVLVEVTTTTSALVKWVSKSAPRVKMYQLQVNCSDDEVLIRMIIPASNKAFVYNN 480
 Db 416 LSQDKIVVAEATSTALLKFNFORNIPGIRMFQIQVNGTYDITLVVYRMIPTTSKTELAVN 475
 QY 481 LVSGTGYDLCLVAMWDDTATLTLATNIVGCAQFFTRADYPOCOSMHSOILGGTMIIVIG 540
 Db 476 LAAAGTMYDLCLVLAIVYDDGITSLATRVRVGCIFTEQDYVRFCHFMQSQGLGTTMIIGG 535
 QY 541 IIVATLLVFIIVLMVRYKVCNHEAPSKMAAAVSNVYSQTNGAQQPPPPSSAPAGAPQGGP 600
 Db 536 IIVASVLFVFIIVLMVRYKVCNNGQHKV-TKVSNIYVSTNGAQ-----IOGCSVTLFQSVS 590
 QY 601 KVVVRNELLDFEFTASLARASDSS---SSSLGSGEAAAGLGRAPWRIIPPS----- 645
 Db 591 KQAVGHE---ENAOCKKATSDNVIQSSEFCSSODSSTTTSA---LPPSWTSTVTSQKOK 644
 QY 646 -APRPKPSIDRLMGAFASLDLKSQRKEE-----LLDSRTPAGRCAGTARSAGHSDREPLG 700
 Db 645 RKTGTPKPEQNEAVTNVESQNTNRNNSNTALQLASRPDPSVTEGTSKRAHIKPSKFTI 704

Query Match 44.5%; Score 1820.5; DB 23; Length 719;
 Best Local Similarity 52.1%; Pred. No. 4.3e-130;
 Matches 371; Conservative 114; Mismatches 190; Indels 37; Gaps 11;

QY 1 METLLGGLLAFGMAFVVDACPKYCVCONLSESLGTLCPKGLLVFPDIDRRVVELRLG 60
 Db 1 MEKILFYLFLLGIA-VKQAICPKRCVCOILSPNLATLCAKGLLFPVPPNDRRIVVELRLA 59

QY 61 GNFIHISRODPANMTGLVLDLTLRSNTIISHIQPFSLDLESRLSHLDNSRLSLGDETL 120
 Db 60 DNFTVNIKRKDFANMTSLVLDLTLRSNTISFITPHAFADLRNLRALHLNSRLTKITNDMF 119

QY 121 RGLVNLQHLIYNNQLGGIADAEFDLLLEDDLSYNNLHGLPWDSVRRMNLHLQSL 180
 Db 120 SGLSNLHLIILNNQLTLISSTAADD-VFALEELDLSYNNLETIPWDAVEKVVSLHTLSL 178

QY 181 DNNLLDHTAECTFADLQKARLDLTSNRLQKLPDPIFARSOASALATTAPEAPPLSFG 240
 Db 179 DNNMIDNIPKGTFSHLHKWRLDVTSNKIQKLPDPLFQRAQVLAATSGIISPSFALSFG 238

QY 241 GNPLHNCCELLWLRRLERDDLETCGSPGGLKGRYFVHVREEFVCEPLLIQTHTKLLV 300
 Db 239 GNPLHNCCELLWLRRLSREDDLETCASPLLTGRYFWSIPEEFLCEPLLIHTRHMERV 298

QY 301 LEGQAATLKCKAIGDPSPLIHWAPDDRLVGNSSRTAVYDNGTLDIFITTSQDSGAFICI 360
 Db 299 LEGQRATLRCARGDPEPAIHWSPEGLIENATRSLVYDNGTLDLITTKVDTGAFICI 358

QY 361 AANAAGEATMVVEISIVOLPHLSNSTRTAPPKSRSLSDITGSSKTSRGGGGGGEPPKS 420
 Db 359 ASNPAGEATQIVDLHIKPLHLLNSTNHIHEPDPGSSDIISTKSGSNWSSNG---DTK 415

QY 421 PPERAVLYSEVTTTSALVKVSKSAPRVKMYQLQYNCSDDEVLVYRMIPIASNKAFVNN 480
 Db 416 LSQDKLVVAEATSSALLKFNQFQNIPIGRWFQIQYNGYDLDLTVYRMIPIPTSKFLVNN 475

QY 481 LVSGTGYDLCVLAWDDTATLITATNIVGCAQFFKADYPCQSMHSQLGTMILLVGG 540
 Db 476 LAAGTMYDLCVLAIVYDGGITSLTATRVVGCIOFTTEQDYVRCHEFMQSOPLGGMIIIGG 535

QY 541 IIVATLLVFIIVLVYKVCNHEAPSKMAAAVSNVYVYVYVYVYVYVYVYVYVYVYVYV 600
 Db 536 IIVASVLFVFIILMRYKVCNNGQHKV-TKVSNNVYVYVYVYVYVYVYVYVYVYVYVY 590

QY 601 KVVVRLNELLDTASLARASDSS---SSSLGSGEAGLGRAPWRIPPS-----645
 Db 591 KQAVGHE--ENAOCKKATSDNVIQSSETCSSQSDSSTTSA---LPPSWTSSTSVSQKQ 644

QY 646 -APRPKPSLDRLMGAFASLDLKSORKEE---LLDSRTPAGRAG-TSARGH 691
 Db 645 RKTGTKPSTEQNEAVTNVESQNTNRNNSNTALQASRPDPSVTEGFTSKRAH 696

RESULT 7
 AAG67512
 ID AAG67512 standard; Protein; 771 AA.
 XX
 XX AAG67512;
 XX
 DT 26-NOV-2001 (first entry)
 XX
 DE Amino acid sequence of a human secreted polypeptide.
 KW Human; secreted polypeptide; nervous disease; muscular disease; tumour;
 KW gastrointestinal ulceration; spinal cord disease; trachea disease;
 KW thyroid gland disease; ovary disease; prostate disease; heart disease;
 KW renal gland disease; small intestine disease; thymus disease;
 KW lymph node disease; muscular system disease; colon disease;
 KW lipase deficiency; cystic fibrosis; pancreatitis; clot formation;
 KW myocardial infarction; angioplasty; liver disease; coagulation disorder;
 KW microbial disease; immune disorder; inflammation; transplant rejection;
 KW bone thickness; bone density; ferroxidase loss; apoptosis;
 KW vascular smooth cell proliferation; vaccine.

QY 701 PPAARARLLPLLEGKAKRSHSFDMGDFAAAAGVWPGGYSPRKVSNITWKBSLWN 760
 Db 705 LPAERSGARHKYSLANGELKEYC-----YINGPNTCGLFPKRSMN 746

QY 761 GMLLPEESD 770
 Db 747 VMFIQDCSD 756

RESULT 6
 ABP43533
 ID ABP43533 standard; Protein; 719 AA.
 XX
 AC ABP43533;
 XX
 DT 08-AUG-2002 (first entry)
 XX
 DE Human secreted protein (SECP) 57.
 XX
 KW Human; secreted protein; SECP; SECP expression; gene therapy;
 KW protein therapy; immune system disorders; AIDS; thymic hypoplasia;
 KW anemia; asthma; Crohn's disease; neurological disorder; epilepsy;
 KW Huntington's disease; dementia; Parkinson's disease; Down's syndrome;
 KW developmental disorder; cell proliferative disorder; cancer.
 XX
 OS Homo sapiens.
 XX
 PN WO200226982-A2.
 XX
 PD 04-APR-2002.
 XX
 PF (INCY-) INCYTE GENOMICS INC.
 XX
 PR 29-SEP-2000; 2000US-236869P.
 PR 11-OCT-2000; 2000US-239812P.
 PR 12-OCT-2000; 2000US-240108P.
 PR 17-OCT-2000; 2000US-241282P.
 PR 20-OCT-2000; 2000US-242218P.
 XX
 PA Yue H., Tang YT, Nguyen DB, Yao MG, Xu Y, Tribouley CM;
 PI Sanjanwala MS, Walia NK, Baughn MR, Sapperstein SK, Lal P;
 PI Thornton M, Gandhi AR, Ramkumar J, Elliott VS, Arvizu C;
 PI Thangavelu K, Gietzen KJ, Ding L, Au-young J, Tran B, Pollicky JL;
 PI Lee S, Lu DAM, Burford N, Warren BA, Gururajan R, Duggan BM;
 PI Honchell CD, Hafalia AJA;
 XX
 WP1: 2002-394239/42.
 DR N-PSDB; ABN99416.
 DR
 XX
 PT New human secreted proteins, useful for diagnosing, treating or
 PT preventing immune system disorders (e.g. Crohn's disease), neurological
 PT disorders (e.g. Parkinson's disease), or cell proliferative disorders
 PT (e.g. cancers)
 XX
 PS Claim 1; Page 188-199; 238pp; English.
 XX
 CC The invention comprises the amino acid and coding sequences of human
 CC secreted proteins (SECP). The SECP DNA and amino acid sequences of the
 CC invention are useful for treating/preventing disorders associated with
 CC increased or elevated expression of SECP. The SECP DNA and protein
 CC sequences are specifically useful for treating/preventing (i.e. gene
 CC therapy and protein therapy): immune system disorders (e.g. AIDS, thymic
 CC hypoplasia, anaemia, asthma or Crohn's disease); neurological disorders
 CC (e.g. epilepsy, Huntington's disease, dementia or Parkinson's disease);
 CC developmental disorders (e.g. Down's syndrome); and cell proliferative
 CC disorders (e.g. cancer). The proteins ABP43477 - ABP43543 represent the
 CC human secreted proteins (SECP) of the invention.
 XX
 SQ Sequence 719 AA;

XX OS Homo sapiens.
 XX PN WO200166690-A2.
 XX PD 13-SEP-2001.
 XX PF 05-MAR-2001; 2001WO-US071143.
 XX PR 06-MAR-2000; 2000US-01871107.
 PR 13-MAR-2000; 2000US-0188916.
 PR 03-OCT-2000; 2000US-0236874.
 PR 03-OCT-2000; 2000US-0237846.
 XX PA (SMIK) SMITHKLINE BEECHAM CORP.
 PA (SMIK) SMITHKLINE BEECHAM PLC.
 XX PI Agarwal P, Murdoch PR, Rizvi SK, Smith RF, Xiang Z;
 XX DR WPI; 2001-570768/64.
 DR N-PSDB; AAH78204.
 XX PT Novel isolated secreted polypeptide useful for treating nervous and
 PT muscular diseases, gastrointestinal ulceration, coagulation and immune
 PT disorders, microbial diseases, inflammation and transplant rejection -
 XX PS Claim 1; Page 69-72; 102pp; English.
 CC The present sequence represents a human secreted polypeptide. The
 CC secreted polypeptides and polynucleotides are useful for treating
 CC nervous and muscular diseases, for inhibiting tumour formation and
 CC metastasis, for treating gastrointestinal ulceration, for preventing
 CC and treating diseases in spinal cord, thyroid gland, ovary, prostate,
 CC renal gland, small intestine, heart, trachea, thymus, lymph node,
 CC muscular system and colon, for treating lipase deficiency in cystic
 CC fibrosis and pancreatitis, for treating undesirable clot formation
 CC such as myocardial infarction, during angioplasty and all surgical
 CC procedures that require decreased blood clot formation, for treating
 CC liver diseases, coagulation disorders and microbial diseases, for
 CC treating immune disorders, for treating inflammation and transplant
 CC rejection, for enhancing bone thickness and increasing bone density,
 CC for reducing the loss of essential ferroxidases, for suppressing
 CC apoptosis, and for regulating vascular smooth cell proliferation. They
 CC may also be used as vaccines.
 XX SQ Sequence. 771 AA;
 Query Match 41.6%; Score 1703.5; DB 22; Length 771;
 Best Local Similarity 46.5%; Pred. No. 4e-121;
 Matches 376; Conservative 110; Mismatches 213; Indels 109; Gaps 18;
 QY 21 CPKVCVQNLSESLGTLCPKSLGFLVPPDIDRRTVELRGGNFIIHSRDFANMTGLVD 80
 Db 34 CPGRCICQNVAPFLMLCAKTLGFLVPPDIDRRTVELRGGNFIIHSRDFANMTGLVD 93
 QY 81 LTLRNTIISHIQFFSFLDLESRLSLDNLPSLGEDTLRGLVNIQHLIVNNOGGIA 140
 Db 94 LTLRNTIQAAGAFADLRAALRALHDSNRLAEVQDQLRGLGNLRLHLGNNQIRRYE 153
 QY 141 DEAFEDFLTLDELDLUSYNNLHGLPWSYRVMVNLHQLSLDNLHDLHIAEGTFADLQKLA 200
 Db 154 SAAFDAPFLSTVEDLDSYNNLEALPWEAVQMVNLTLDHNLIDHIAEGTFVQLKLV 213
 QY 201 RLDTLSNRLOKLPDPIFARSOASALTATPFAP-PLSFSFGGNPLHCNCELLWRLERD 259
 Db 214 RLDMTSNRHLKLPDGLFLRSQGTG-----PKPTPLTVSFGGNPLHCNCELLWRLERTRE 269
 QY 260 DDLETGSPGGLKGRYFVHWRBEFVCEPPLIT-QHTHKLVLLEQQAALKCKAIGDPSF 318
 Db 270 DDLETGATPEHLTDRLFWSIPBEFLCEPPLITRQAGRALVVEGQAVSRCAVGDPEP 329
 QY 319 LIHWAPDRLVGNSSRTAVYDNGTLDFITTSQDSGAFPTCAANAAGBATAVVEYSIVO 378

Db 330 VVHWVAPDGRLLGNSRTRVRGDTLDVTIITLDRSGTFTFCIASNAAGBATAVVEYSIVO 389
 QY 379 LPHLSNSTRTAPPKSRSLSDIITGSSKTSRGGGGGEGPEPKPPPERAVLSEVTTTSALV 438
 Db 390 LFLM-----APPAAPPLTEPGSSDIATFGRPGAN--DSAAERRLVAEELTSNSVLI 440
 QY 439 KWSVSKSAPRVKMYQLOYNCSDDVLIYRMIKASNAKAFVNNLVSGTGTDLGVLANWDDT 498
 Db 441 RWPQRVPEGIRMYQVYNSVDDSLVYRMIPESTQTFVNLNLAAGRAYDLCVLAIVYDDG 500
 QY 499 ATTLTATNIVGCAQPFKADYQOCOSMHSQIILGMMILVIGGIIIVATLVFVILMVRYK 558
 Db 501 ATALPATRVVGVVQVFTTAGDPAPCPRAHFLGGMIIAIGGVIVASVLIIVLMLIRYK 560
 QY 559 V-----CNHEAPSKMAAAVSNVYQTNQAQPPPPPSAPAGAPP-----QGPP 600
 Db 561 VYGDGDSRRVKGSRSLPRVSHVCSQTNGA--GTGAAQAPALPAQDHYEALREVESQAAP 617
 QY 601 KVVVRNELLDFATSLARASDSSSSLSGSGEAGLGRAPWRIPPSAPRKPSPDLRMLMGAF 660
 Db 618 AVAVE-----AKAMEAETASA-----EPEVVLGRSLGGS 646
 QY 661 A-SLDLKSORKEELDSRTPAGRGAGTSARGHHSREPLLGPPAARARSILPLP----LE 715
 Db 647 ATSLCLLPSEETSGEESR-----AAVGRPRSRSGALEPPTSAPPTLALVPGGAAA 697
 QY 716 GKAKRSHSDFMDGFAAAAAGVYVGGYSPRKRVSNIWTKRSLSVNGMLLPFEESDLVGAR 775
 Db 698 PRPQQRYSFD-CDY-----GALFQSHSYPRRARR--TKRHRST-----PHLDGAGGGAA 743
 QY 776 G-----TFGSSEVMWMESTV 789
 Db 744 GEDGDLGLGSARACLAFTSTENWLESTV 771
 RESULT 8
 AAE23980
 ID AAE23980 standard; Protein; 635 AA.
 AC AAE23980;
 XX 23-SEP-2002 (first entry)
 XX Human LP220 secreted protein.
 KW Human; secreted protein; atherosclerosis; Alzheimer's disease; LP220;
 KW diabetic retinopathy; severe combined immunodeficiency; pancreatitis;
 KW rheumatoid arthritis; colorectal adenoma; haemolytic anaemia; cancer;
 KW reperfusion injury; arteriosclerosis; wound healing; transgenic animal;
 KW gene therapy; neoplasm; transgenic; psoriasis; ischaemia; carcinoma;
 KW chromosome 11q13.
 XX Homo sapiens.
 OS
 XX
 FH Key Location/Qualifiers
 FT Peptide 1..16
 FT /label= Signal_peptide
 FT Protein 17..635
 FT /note= "Mature human LP220 secreted protein"
 XX WO200226801-A2.
 XX
 XX PD 04-APR-2002.
 XX PF 14-SEP-2001; 2001WO-US26026.
 XX PR 28-SEP-2000; 2000US-236088P.
 XX PA (ELIL) LILLY & CO ELI.
 XX PI Su EW, Wang H;
 XX WPI; 2002-471259/50.
 DR

CC at ftp.wipo.int/pub/published_pct_sequences.
 XX SQ Sequence 322 AA;
 Query Match 40.4%; Score 1656; DB 22; Length 322;
 Best Local Similarity 100.0%; Pred. No. 5e-118;
 Matches 322; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 468 MIPASNKAFVNNLVSGTYDLCVLAAMDDDTATLTATNIVGCAOFFTKADYPOCOSMHS 527
 Db 1 MIPASNKAFVNNLVSGTYDLCVLAAMDDDTATLTATNIVGCAOFFTKADYPOCOSMHS 60

QY 528 QILGTMILVIGGIIVAILLPIVILMVRKYCNHEAPSKMAAAVSNVYQINGAQPPPP 587
 Db 61 QILGTMILVIGGIIVAILLPIVILMVRKYCNHEAPSKMAAAVSNVYQINGAQPPPP 120

QY 588 SSAPAGAPQGGPKVVRNELLDFASLARASDSSSSSLGSGEAGLGRAPWRIPPSAP 647
 Db 121 SSAPAGAPQGGPKVVRNELLDFASLARASDSSSSSLGSGEAGLGRAPWRIPPSAP 180

QY 648 RPKPSLDRLMGAFASLDLKSQRKELLDLSDRTPAGRGAGTSARGHSDREPLLGPPAARAR 707
 Db 181 RPKPSLDRLMGAFASLDLKSQRKELLDLSDRTPAGRGAGTSARGHSDREPLLGPPAARAR 240

QY 708 SLLPLPLEGKAKRSHSFDMGDFAAAAAGGVVPGGYSPPRKVSNITKRSLSVNGMLLPE 767
 Db 241 SLLPLPLEGKAKRSHSFDMGDFAAAAAGGVVPGGYSPPRKVSNITKRSLSVNGMLLPE 300

QY 768 ESDLVGARTFGSSSEWMESTV 789
 Db 301 ESDLVGARTFGSSSEWMESTV 322

RESULT 11
 AAB84469
 ID AAB84469 standard; Protein; 628 AA.
 XX AC AAB84469;
 XX DT 05-SEP-2001 (first entry)
 XX DE Amino acid sequence of an interferon omega-1 like protein NOV2.
 XX KW Interferon omega-1 like protein; Nov2; membrane bound protein;
 CC secreted protein; spermatogenesis; male infertility; neoplasia;

CC identifying agents (agonists and antagonists) that bind to them. Cells
 CC expressing the proteins are useful for identifying a therapeutic agent
 CC for use in treatment of a pathology related to aberrant expression or
 CC physiological interactions of the polypeptide. Vectors comprising
 CC the nucleic acids encoding the polypeptides and cells genetically
 CC engineered to express them are also useful for producing the proteins.
 CC The proteins are useful in genetic vaccination, testing and
 CC therapy, and can be used as nutritional supplements. They may be used to
 CC increase stem cell proliferation; to regulate haematopoiesis; and in
 CC bone, cartilage, tendon and/or nerve tissue growth or regeneration;
 CC immune suppression and/or stimulation; as anti-inflammatory agents; and
 CC in treatment of leukaemias. AAU29510-AAU33304 represent the amino acid
 CC sequences of novel human secreted proteins of the invention.
 XX SQ Sequence 636 AA;
 Query Match 38.8%; Score 1586.5; DB 22; Length 636;
 Best Local Similarity 53.5%; Pred. No. 2.6e-112;
 Matches 315; Conservative 88; Mismatches 171; Indels 15; Gaps 6;

QY 16 AVVDACPKYCVYQNLSESLGTLCPKSGLLFVPPDIDRRITVELRLGNGFIHISRODFANM 75
 Db 23 ATPSPCRRRCQTCQSLPLSVLPCGAGLLFVPPSLDRRAAELRADNFIAVRRRDLANM 82

QY 76 TGLVDLTLNRNTHIHIQPFSLDLESLSLHSDNRLPSLGEDTLRGLVNLQHLIVNNQ 135
 Db 83 TGLLHLSLRNTHIIRVAAGAFADLRALRALHLDGNRLTSLGEGQLRGLVNLRLHLSNNO 142

QY 136 LGGIADAEFEDLLELDSYNNLHGLPWDSVPRMNLHOLSDHLDLHIAEGTFAD 195
 Db 143 LAALAAAGALDDCAETLELDSYNNLEQLPWEALGRGNVNTLGLDHLNLLASVPAGAFSP 202

QY 196 LOKLARLDLTSNRLOKLPDPPIFARSQASALTATPAPPPLS---FSFGNPLHNCLELW 252
 Db 203 LHKLARLDLTSNRNLTITPPDPLFSR--LPLLAREPGSPASALVLAFCGNPLHNCLELW 259

QY 253 LRRLEDDDLTCGSPGGKGRYFVHWREBEFVCEPPLITQHTKLLVLEGGQAATLKCKA 312
 Db 260 LRRLAREDDLEACASPPALGGRYFWAVGEBEVEFVCEPVPVTHRSPLVAPGAPRALRCRA 319

QY 313 IGDPSLIHWAPDDRLVGNSSRTAVYDNGTLDIFITTSODSGAFTCIAANNAAGATAMV 372
 Db 320 VGDPEPRVVRVSPQGRLLGNSSRARAFNGTLELLELVEFDGDIIFTCIAANNAAGATAA 379

QY 373 EVSI--VOLPHLSNSTRTAPPKRLSD-ITGSSKTSRGGGGGGGPKSPPEPRAVLVS 429
 Db 380 ELTVGPPPPQLANSTSCDPPGRCGLPDALFPPSAASASAKVADTG---PPTDRGVQVT 435

QY 430 EVTTTSALVKWSVSKSAPRVKMYQLQYNCSDDEVLIRMI PASNKAFVNNLVSGTYDYL 489
 Db 436 EHGATAALVQWPQRPPIGIRMYQIQYNSADDLIVYRMI PAESRSFLLDLASCRTYDL 495

QY 490 CVLAWMDDTATLTATNIVGCAOFFTKADYPOCOSMHSQILGTMILVIGGIIVAILL 549
 Db 496 CVLAVYEDSATGLTATREVGCARFSTEPSLRPCGAPHAPFLGGTMIILAGGVIVASVLVF 555

QY 550 IIVLMVRKYCNHEAP--SKMAAAVSNVYQINGAQPPPPSAPAGAPP 596
 Db 556 IFVLLMRYKVGQPPGKAKIPAPVSSVCSQTNALGFTPTTAPPAPPEP 604

RESULT 11
 AAB84469
 ID AAB84469 standard; Protein; 628 AA.
 XX AC AAB84469;
 XX DT 05-SEP-2001 (first entry)
 XX DE Amino acid sequence of an interferon omega-1 like protein NOV2.
 XX KW Interferon omega-1 like protein; Nov2; membrane bound protein;
 CC secreted protein; spermatogenesis; male infertility; neoplasia;

Best Local Similarity 52.9%; Pred. No. 9.2e-111; Mismatches 175; Indels 14; Gaps 5; Matches 311; Conservative 88;

QY 16 AVVDACPKYCVCONLSESI... 75
 Db 23 ATPSPCPRRCRCOTQSL... 82
 QY 76 TGLVDLTLSSNTIRHVA... 135
 Db 83 TGLLHLSLSRNTIRHVA... 142
 QY 136 LGGIADAEFEDLLTLE... 195
 Db 143 LAALAGALDDCAETL... 202
 QY 196 LOKLARLDLTSNRLQ... 252
 Db 203 LHKLARLDLTSNRL... 259
 QY 253 LRRLERDDLETCGSG... 312
 Db 260 LRRLAREDDLEACAP... 319
 QY 313 IGDPSLIIHWVADDR... 372
 Db 320 VGDPEPRVRSVQGR... 379
 QY 373 EVSI--VQLPHLSN... 430
 Db 380 ELIVGPPPPQLANS... 435
 QY 431 VTTTSALVKWSKSA... 490
 Db 436 HGATAALVQWPDQ... 495
 QY 491 VLAMWDDTATLTAT... 550
 Db 496 VLAVGESATGLTAT... 555
 QY 551 VILMRYKVCNHEAP... 596
 Db 556 FVLLMRYKVGQPPG... 603

RESULT 14
 ABG34079
 ID ABG34079 standard; Protein; 627 AA.
 AC ABG34079;
 XX
 DT 15-JUL-2002 (first entry).
 DE Human Pro peptide #51.
 DE Human; PRO; secreted protein; transmembrane protein;
 KW genetic disorder; tumour; cancer.
 KW Homo sapiens.
 OS WO200224888-A2.
 FN
 XX
 XX 28-MAR-2002.
 XX
 XX 29-AUG-2001; 2001WO-US27099.
 XX
 XX 01-SEP-2000; 2000US-229896P.
 PR 05-SEP-2000; 2000US-230621P.
 PR 22-SEP-2000; 2000US-235147P.
 PR 10-NOV-2000; 2000WO-US30873.
 PR 12-JAN-2001; 2001US-261878P.
 PR 16-JAN-2001; 2001US-261910P.
 PR 16-JAN-2001; 2001US-261939P.
 PR 16-JAN-2001; 2001US-262150P.

25-JAN-2001; 2001US-264395P.
 02-FEB-2001; 2001US-266421P.
 09-FEB-2001; 2001US-267623P.
 28-FEB-2001; 2001WO-US06520.
 09-MAR-2001; 2001US-274399P.
 03-APR-2001; 2001US-280982P.
 04-APR-2001; 2001US-282129P.
 04-APR-2001; 2001US-282199P.
 09-MAY-2001; 2001US-290589P.
 25-MAY-2001; 2001WO-US17092.
 01-JUN-2001; 2001WO-US17800.
 20-JUN-2001; 2001WO-US19692.
 29-JUN-2001; 2001WO-US21066.
 09-JUL-2001; 2001WO-US21735.
 (GETH) GENENTECH INC.
 Baker KP, Eaton DL, Filvaroff E, Goddard A, Grimaldi JC,
 Gurney AL, Smith V, Stephan J, Watanabe CK, Wood WI, Zhang Z;
 Fong S;
 WPI: 2002-362426/39.
 N-FSDB; ABK70011.
 New PRO polypeptides and polynucleotides encoding the polypeptides,
 useful in gene therapy, chromosome identification, tissue typing, or
 for genetic analysis of individuals with genetic disorders -
 Claim 11; Figure 102; 218pp; English.
 This invention relates to the cDNA and protein sequences of novel
 secreted and transmembrane polypeptides PRO polypeptides. The
 invention also comprises a method for producing the proteins of the
 invention by recombinant means and antibodies specific for the protein
 of the invention. The antibody may be used for detecting the PRO
 proteins of the invention and may be used to modify their activity.
 Polynucleotides may be used as hybridisation probes for a cDNA library
 to isolate the full-length PRO cDNA or to isolate other cDNAs, to
 PRO and for genetic analysis of individuals with genetic disorders, in
 assays to identify other proteins or molecules involved in binding
 reaction, to generate transgenic animals or knock-out animals which
 turn are useful in the development and screening of therapeutically
 CC useful reagents, for chromosome identification and tissue typing. The
 CC PRO polypeptides are useful in gene therapy, and as molecular weight
 CC markers for protein electrophoresis purposes. The sequences may
 CC also be used to detect overexpression on PRO polypeptides in cancerous
 CC tumours and for screening for differentially expressed genes using
 CC microarray technology. The present sequence represents a human PRO
 CC protein of the invention.
 XX
 SQ Sequence 627 AA
 Query Match 38.2%; Score 1563.5; DB 23; Length 627;
 Best Local Similarity 53.1%; Pred. No. 1.4e-110;
 Matches 312; Conservative 88; Mismatches 173; Indels 15; Gaps 6;
 QY 16 AVVDACPKYCVCONLSESI... 75
 Db 23 ATPSPCPRRCRCOTQSL... 82
 QY 76 TGLVDLTLSSNTIRHVA... 135
 Db 83 TGLLHLSLSRNTIRHVA... 142
 QY 136 LGGIADAEFEDLLTLE... 195
 Db 143 LAALAGALDDCAETL... 202
 QY 196 LOKLARLDLTSNRLQ... 252
 Db 202 LHKLARLDLTSNRL... 259

QY 253 LRLRLRDDLETCGSPGLKGRYFHWVREERFVCEPPLITQHTKLLVLEGOAATLKCKA 312
 Db LRLRLRDDLEACASPPALGGRYFVAVGEEFVCEPPLVTHRSPLVAVPAGRPALRCRA 318
 QY 313 IGDPSPLHWTAPDRLVNSRRTAVYDNGTLDIFITTSQDSGAFCTIAANAAGATAMV 372
 Db VGDPEPRVWSPQGRLLGNSRRARAFNGTLELVLVTEPGDGIFFCIAANAGATAV 378
 QY 373 EVSI--VQLPHLSNSTRTAPPKSLSDITGSSKTSRGGGSGGEGEPPKSPERAVLYSE 430
 Db ELTVGPPPPOLANSTSCDPPRDGDPDALTPPSAASAKVADTG---PPTDRGVQVTE 434
 QY 431 VTTISALVKWSVSKAPRVMYQVYQYNSDDEVLVIRMPASNKAFVNNVNLVSGTCYDLC 490
 Db HGATAALVQWPPQRPPIGIRMYQIQYNSADDILVIRMPAESRFLTLTLASGRTYDLC 494
 QY 491 VLAMDDTATLTATNIVGCAOFFTKADYPOCOSMHSOILGQTMILVIGGIIVATLLVFI 550
 Db VLAVYEDSATGLTATRPVGCARFSTEPALRPCGAPHAFPLGGTMIALGGVIVASVLVFI 554
 QY 551 VILMRYKVCNHEAP--SKMAAAVSNVYSQTNAGQPPPPSSAPAGAPP 596
 Db 555 FVLLMRYKVGQPPGKAKIPAPVSSVCSQTNALGPTPTAPPAPPEP 602

RESULT 15
 AAEL17484
 ID AAEL17484 standard; Protein; 551 AA.
 AC AAEL17484;
 XX

DT 22-APR-2002 (first entry)
 XX Human leucine-rich repeat-8 (ZLRR8) protein #2.
 KW Human; leucine-rich repeat-8; ZLRR8; cytosstatic; gene therapy; leukaemia;
 KW endometrial adenocarcinoma; renal cell; colon; prostate; kidney tumour; lung;
 KW cell growth disorder; glioblastoma; neuroblastoma; kidney tumour; lung;
 KW mammary gland; germ cell; bladder; oesophagus; pancreas; animal feed;
 KW rhabdomyosarcoma; genitourinary tract; chromosome 11q13.
 XX Homo sapiens.

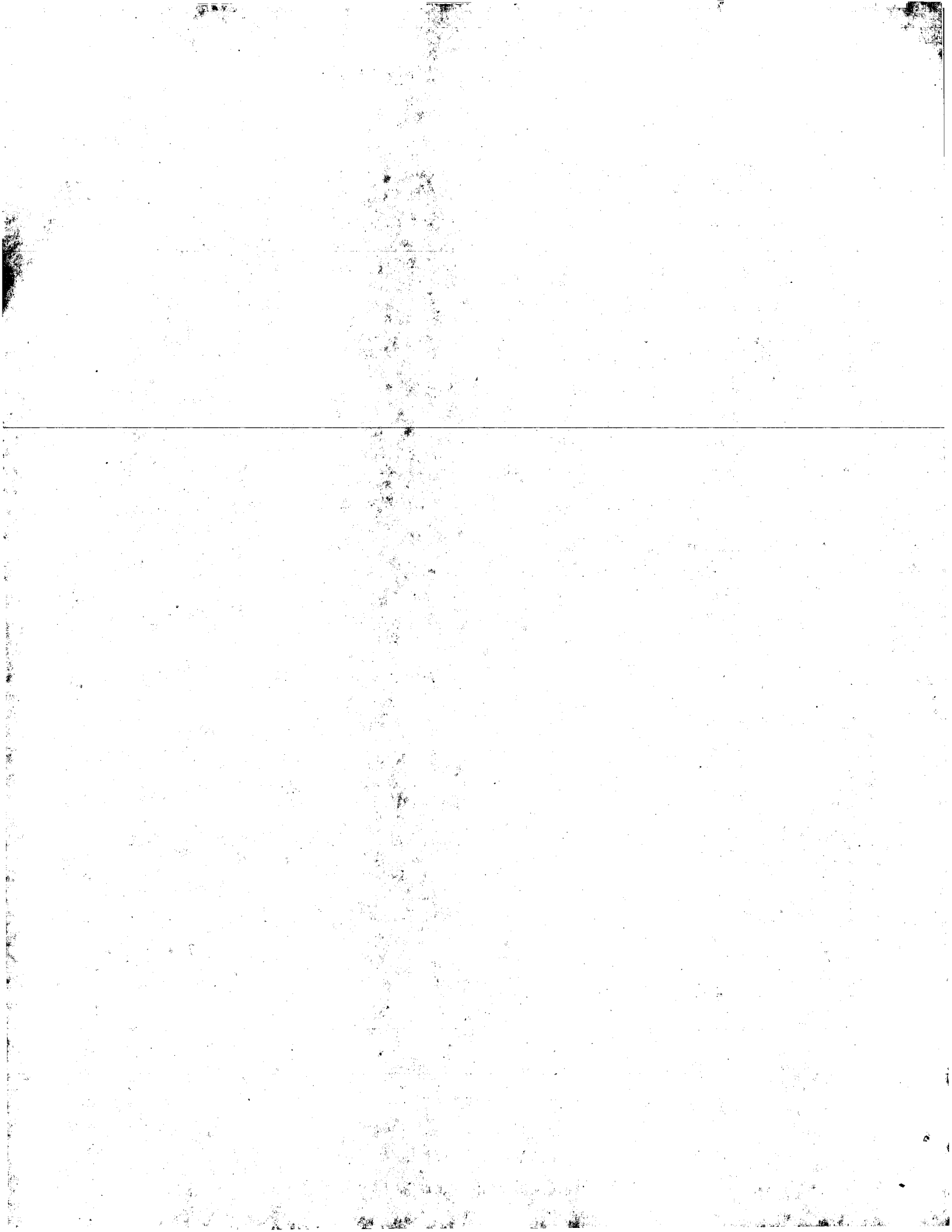
XX Key Location/Qualifiers
 FH Peptide 1..18
 FT /label= Signal_peptide
 FT Protein 19..551
 FT /label= Mature_Zlrr8_protein
 XX WO200202604-A2.
 XX 10-JAN-2002.
 XX 02-JUL-2001; 2001WO-US20999.
 XX 30-JUN-2000; 2000US-215446P.
 XX (ZYMO) ZYMOGENETICS INC.
 XX Thayer EC, Sheppard PO, Presnell SR;
 XX WPI; 2002-154725/20.
 XX N-PSDB; AAD28124.
 XX New leucine-rich repeat proteins and polynucleotides, useful for
 XX diagnosing and treating disorders related to abnormal cell growth e.g.
 XX retinoblastoma, renal cell, endometrial adenocarcinoma, leukemia,
 XX kidney and lung tumors
 XX Claim 10; Page 67-69; 82pp; English.
 XX The invention relates to an isolated polypeptide comprising leucine-rich
 XX repeat proteins Zlrr7, Zlrr8, Zlrr9 and spliced variants of Zlrr8, Zlrr9.

CC Zlrr7, Zlrr8, and Zlrr9 proteins are useful in directing the secretion of
 CC proteins of interest from a host cell and to monitor the secretion of
 CC proteins in general from cells and tissues. The Zlrr DNA and proteins
 CC are useful in diagnosing and treating disorders related to abnormal cell
 CC growth, including retinoblastoma, renal cell adenocarcinoma, endometrial
 CC adenocarcinoma, glioblastoma, neuroblastoma, B-cell lymphocytic leukaemia,
 CC kidney tumours, germ cell tumours, lung large cell carcinoma, mammary,
 CC colon adenocarcinoma, genitourinary tract transitional cell tumours,
 CC rhabdomyosarcoma, lung tumour, bladder tumour, oesophagus, pancreas
 CC and prostate adenocarcinoma. Zlrr protein is useful for identifying
 CC agonists and antagonists of the polypeptide, for drug design, to screen
 CC for cell metabolism effecting receptors, for analysis of cell phenotype,
 CC and as animal feed supplement and cell culture components. Zlrr DNA is
 CC also useful in gene therapy. The present sequence is human Zlrr8 protein.
 CC Zlrr8 gene is located on chromosome 11q13.
 XX

QY Sequence 551 AA;
 Query Match 37.8%; Score 1545.5; DB 23; Length 551;
 Best Local Similarity 57.6%; Pred. No. 2.8e-109;
 Matches 311; Conservative 70; Mismatches 140; Indels 19; Gaps 6;

QY 8 LLAGMAFVVDACPKYCVCONLSESLGTLCPKGLLFPVPPDIDRRTVELRLLGGNFIHI 67
 Db 9 LLASGAA-----ACPLPCVCONLSESLTLCARHLLFPVPPVDRRTVELRLADNFIOAL 63
 QY 68 SRQDFANMTGLVDLTLRSNTISHIQPFSLDLESLSRLHLDNSRLPSLGEDTLRGLVNLQ 127
 Db 64 GPPDFRNTWGLVDLTLRSNAITRIGARAFGDLSELSRLHLDGNRLVELGTSLRGPVNLQ 123
 QY 128 HLIWNNQGGIADAEFEDFLLEDLDSYNNLHGLPMDSVRRMVLHQLSLDHLNLDH 187
 Db 124 HLLLSGNQLGRIAPCAFDDFLESLEDLDSYNNLRQVWPWAGIGAMPALHTLNLNLDIDA 183
 QY 188 IAEFTFADLQKLRLDILTSNRLQKLPDPIPARSOASALTATPPAPPPLSFSGGNPLHCN 247
 Db 184 LPPGAFAGLQGLSRDLTSNRLATLAPPLFSRGRDA--EASP--APLVLFSGPNLHCN 239
 QY 248 CELLMRLERDDDLLETGCSFGGLKGRYFHWVREERFVCEPPLITQHTKLLVLEGOAAT 307
 Db 240 CELLMRLERLARPDDLETGCSFGGLKGRYFHWVREERFVCEPPLITQHTKLLVLEGOAAT 299
 QY 308 LKCKAIGDPSPLIHWVAPDDRLVGNSSRTAVYDNGTLDIFITTSQDSGAFCTIAANAAGE 367
 Db 300 LRCRALGDPAPTMMHWGPDRLVGNSSRRARAFNGTLEIGVTGAGDAGGYTCIATNPAGE 359
 QY 368 ATMVEVSIVOLPHLSNSTRTAPPKSLSDITGSSKTSRGGGSGGEGEPPKSPERAVL 427
 Db 360 ATARVELRVLALPHGGNSSAEGGRPGP--SDIAASARTAAEBEGTLESEP-----AVQ 410
 QY 428 VSEVTTTTSALVKWSVSKAPRVMYQVYQYNSDDEVLVIRMPASNKAFVNNVNLVSGTCY 487
 Db 411 VTEVTATSGLVSWGEPGADPVMWFOIQYNSDETLVIRVFPASSHFLKHLVPGADY 470
 QY 488 DCLVAMWDDTA-TTLTATNIVGCAOFFTKADYPOCOSMHSOILGQTMILVIGGIIVATL 546
 Db 471 DLCLLALSPAAGPSDLTATLTLGCAHFSTLPSPLCHALQAHVLTGGLTTLVAVGGVLAAL 530

Search completed: June 8, 2003, 21:28:18
 Job time : 68 secs



GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 8, 2003, 21:27:43 ; Search time 36 Seconds
(without alignments)
644.852 Million cell updates/sec

Title: US-09-831-846-2
Perfect score: 4094
Sequence: 1 METLLGGLLAFGMFAVDA.....DLVGARGTGSSEWNESTV 789

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : Issued Patents AA.*
1: /cgn2_6/prodata/2/aaa/5A_COMB.pep.*
2: /cgn2_6/prodata/2/aaa/5B_COMB.pep.*
3: /cgn2_6/prodata/2/aaa/6A_COMB.pep.*
4: /cgn2_6/prodata/2/aaa/6B_COMB.pep.*
5: /cgn2_6/prodata/2/aaa/PCTUS_COMB.pep.*
6: /cgn2_6/prodata/2/aaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Table with columns: Result No., Score, Query Match, Length, DB ID, Description. Contains 27 rows of search results.

Table with columns: 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 44, 45. Contains sequence alignment data.

ALIGNMENTS

RESULT 1
US-09-131-648-2
; Sequence 2, Application US/09131648
; Patent No. 6168920
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Yue, Henry
; APPLICANT: Corley, Neil C.
; APPLICANT: Guegler, Karl J.
; APPLICANT: Patterson, Chandra
; TITLE OF INVENTION: EXTRACELLULAR ADHESIVE PROTEINS
; FILE REFERENCE: PF-0576 US
; CURRENT APPLICATION NUMBER: US/09/131,648
; CURRENT FILING DATE: 1998-08-10
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PERL Program
; SEQ ID NO 2
; LENGTH: 708
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: 2687731
US-09-131-648-2

Query Match 9.0%; Score 368; DB 4; Length 708;
Best Local Similarity 22.6%; Pred. No. 3.4e-22;
Matches 154; Conservative 98; Mismatches 238; Indels 192; Gaps 27;
QY 1 METLLGGLLAFGMFAVDA---CPKYCVQ-----NLSESLGTL-CPKGLLVFPP 48
Db 8 IHVLIG--LAITTLVQAVDKKVDKPCRLCTCEIRPWFPTSRISYMEASTVDCNDLGLLTFPA 65
QY 49 DIDRTEVLRLLGNFIHIS-RQDF-ANMTG----- 77
Db 66 REPANTQILLQTNNAKTEYSTDFPVNLTGLDLQNNLLSVTNVKKRMPQLLSVYLEE 125
QY 78 -----LVDLTLRSNTIHSIQFSPFLDLESLRSLHLDSDNRLPSLGEDTURL 121
Db 126 NKLTLPKELKSELNLOELYNHNLLSTISPGAFIGLHNLRLHLNSRNLQMINSKWFD 185
QY 122 GAVNLOHLIVNNQLGGTAEAF-----EDFLTLEDLD-LSY-- 158
Db 186 ALPNEILMIGENPIIRIKDMFKLNLRSUVIAGINLTEIPDNALVGLNLEISISFYD 245
QY 159 NNHLGLPWSVRRMVLHQLSLDHLNLDHIAEGTFA----- 194
Db 246 NRLIKVPHVALQKVVNLFKFLDKPKPIINRIIRGDFSNMLHLKELGINNPELISIDSLAV 305
QY 195 -DLQKLARLDLTSN-RLQKLPDPPIFARSOASALATPFA-----PPL-SF5F 239

Db 306 DNLPLDKTEATNPRLSYIHPNAPFRLPKLESMLNSALYHGHTIESLPLKLEISI 365
 Qy 240 GGNPLHCNCELLWLR-----RLERDDDLLETGSPGLKGRYFVHWREBEF-----VCEP 288
 Db 366 HSNPIRCDVCVIRMMNKNIRMFEPDSL-FCVDPPEFQO---NVRQVHFRDMMIEICLP 421
 Qy 289 PLITQHTHKLLE-QQAATLCKAIGDPSPLIHWAPD-DRLVGN--SSRTAVYDNGTL 344
 Db 422 LIAPESFPNLNVEAGSVYFHCRTAEFQPELIYWIITPSGQKLLNLTLDKFFVHSEGL 481
 Qy 345 DIFITTSODSGAFTCIAANAAGEATAMVEVSIVOLPHLSNSTSRTPPKSRISLDTGSSK 404
 Db 482 DINGVTPKGGGLYTCIATNLV-----ADLKSVMIKVDGSP 518
 Qy 405 TSRGGCGSGGPPKPPPERAVLVSIVITTSALYKWSV-----KSAPRVVMYQLQYNCS 460
 Db 519 QDNNGSLN-----IKIRDIQANSVLVSWKASSKILKSSVKWTF--VKTEN 562
 Qy 461 DEVLIVRMI PASKAFVNVNLSVGTGCDLCLVAMWDDTATLTATNIVGCAQFFTKADYP 520
 Db 563 SHAQASARIPSDVKVYNLTHLPSTYKICI-----DIPTIYQNRKCKVNVYTKGLHP 616
 Qy 521 QCQSMHSQILGTMILVIGII 542
 Db 617 D-QKEYEKNTTTLMACLGELL 637

RESULT 2

US-08-986-485-5
 ; Sequence 5, Application US/08986485
 ; Patent No. 6046030
 ; GENERAL INFORMATION:

; APPLICANT: WU, SHUJIAN
 ; APPLICANT: SWEET, RAYMOND
 ; APPLICANT: TRUNEH, ALEMSEGED
 ; TITLE OF INVENTION: A HUMAN LIG-1 HOMOLOG (HLIG-1)
 ; NUMBER OF SEQUENCES: 8

; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: RATHER & PRESTIA
 ; STREET: P. O. BOX 980
 ; CITY: VALLEY FORGE
 ; STATE: PA
 ; COUNTRY: USA
 ; ZIP: 19482

; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: .Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: DOS
 ; SOFTWARE: FASTSEQ for Windows Version 2.0
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/986,485
 ; FILING DATE: 08-DEC-1997

; CLASSIFICATION:
 ; PRIORITY APPLICATION DATA:
 ; APPLICATION NUMBER: 60/059,448
 ; FILING DATE: 22-SEP-1997
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: PRESTIA, PAUL F
 ; REGISTRATION NUMBER: 23,031
 ; REFERENCE/DOCKET NUMBER: GH-70264
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 610-407-0700
 ; TELEFAX: 610-407-0701
 ; TELEX: 846169

; INFORMATION FOR SEQ ID NO: 5:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1091 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein

US-08-986-485-5

Query Match 9.0%; Score 368; DB 3; Length 1091;
 Best Local Similarity 21.3%; Pred. No. 6.6e-22;
 Matches 205; Conservative 92; Mismatches 326; Indels 340; Gaps 33;

Qy 30 LSESLGTLCPKSKGLL-----FVPPDIDRRTVELRLCGNFIIHISRODFANMTGLVDLTF 82
 Db 188 LRSRLLTLRLSKNRITQLPVKAFKLP-----RLTQLDLNRNRIRLIEGTFQGLDLSLEVL 243
 Qy 83 LSRNTI SHIOPPSFLDLSRSLHLSNLSRSLGDTLRLGVLNQLQHLIWNHNLGGIADE 142
 Db 244 LQRNNISRLTQGAFWGLSKMHLHLEYNLSLVEVNSGSLYGLTALHQLHLSNNSISRIORD 303
 Qy 143 AFEDFLLEDLDSYNNLHGLPMDSVRRMNLHOLSLDHLLDHLAEGTFAADLQKLRAL 202
 Db 304 GW-SFCQKHELHLLSFFNNLTRLDEBSLAELSSLSLRLSHNAISHIAEGAFKGLKSLRVL 362
 Qy 203 DLTSNRLQKLPDPIPARSOASALT-ATPFAPPLS-----PSFGGNPLH- 245
 Db 363 DLDHNEISGTTIEDTSGAFTGLDNLKSLTFLGNKIKSVAKRAFSGLESLEHLNLGENAIRS 422
 Qy 246 -----CNCELLWL-----RRLERDDDLLETGSPGLKGRYFW 277
 Db 423 VOFDAFAKMKLKELYISSESFCLDCQKLPWLPMLMGRMLQAFVTATCAHPESLKGQSI 482
 Qy 278 HVREEFVYCE-----PPLIITQ----- 293
 Db 483 SVLPDSFVCDPFPKPIITQPETTMAVYKDIRFTCSAASSSSSPMTFAWKKDNEVLANA 542
 Qy 294 -----HTHK----- 297
 Db 543 DMENFAHRAQDGEVMEYTTILHLRHVTFEGRYQCIITNHFSTYSHKARLTVNVLPS 602
 Qy 298 -----LLVLEGOAATLKCKAIGDPSPLIHW-----VAPDDRUVGNS 333
 Db 603 FTKIHPDIAIRGTTRLECAATGHPNFQIAWKQGGDTFFPAARERRMHWEMDD----- 656
 Qy 334 SRTAVYDNGTLDLFIITTS---ODSGAFTCIAANAAGEATAMVEVSIVOLPHLSNSTSR 390
 Db 657 -----DVFFITDVKIDDMGVYCTAQSAGSVSANATLVLETPEL-----A 698
 Qy 391 PPKSRLSLDTGSSKTSRGGSGGEGEPPKSPERAVLVSIVITTSALYKWSVSKSAPRYK 450
 Db 699 VELE-----DRVVTVGE---TVAFOCKATGSPPTPRIT 727
 Qy 451 MYLOQVNGSDDEVLIVRMI PASKAFVNVNLSVGTGCDLCLVAMWDDTAT-TLTATNIVG 509
 Db 728 WLKGRPLSLTE--RHHFTP-GNQLLVQNV-----MIDDAGRYTCENSNPLG 772
 Qy 510 CAQFFTKADY---POCQSMHSQILGTMILVIGIIAVTLLVFIILMVRVYKVCNHEAFS 566
 Db 773 TERAHSQLSLPTPGCRK-DGTTVIGIFTIAVVCISVLTSLVWVCIIYOTRKK----- 823
 Qy 567 KWAAAVSNVYQTNGAQPPSSAGAPQGP-----PKVVYRNE----- 607
 Db 824 -----SEEYSVTNTDETIVPDPVPSYLSQGTLSDRQETVVRTEGHOANGHIESNGVC 877
 Qy 608 LLDFTASLARASDSSSSSLGSGEAGLGRAPFWRIPPSAPR-PKPSLDRLMGAF----AS 662
 Db 878 LRD--PSLPEYVDIHSITCRQPKLVGYTRPFWKTERADRATAAPHTTAHSSGSAVCSDCS 935
 Qy 663 LDLKSRKEELLSRTPAGRGAGTSARGHSHSDREPLLGPAPARAKSLLEPLLEGKAKRSH 722
 Db 936 TDTAYHPQVPRDSGOP-GTASSQBLRQHDREYSP-----HHPYSGTADGGH 981
 Qy 723 SEDMGDFAAAAAGGVVPGYSPRKVSNLWTKESLVNGMLLFFESDLVARGTGTGSSS 782
 Db 982 TL-----SGGSLYPSNHD-----RIPLPSLKNKAASADGN-GDSS 1014
 Qy 783 WVM 785
 Db 1015 WTL 1017

RESULT 3
 US-08-986-485-2 ; Sequence 2, Application US/08986485 ; Patent No. 6046030 ; GENERAL INFORMATION ; APPLICANT: WU, SHUJIAN ; APPLICANT: SWEET, RAYMOND ; APPLICANT: TRUNEH, ALEMSEGED ; TITLE OF INVENTION: A HUMAN LIG-1 HOMOLOG (HLIG-1) ; NUMBER OF SEQUENCES: 8 ; CORRESPONDENCE ADDRESS: ADDRESSEE: RATNER & PRESTIA ; STREET: P.O. BOX 980 ; CITY: VALLEY FORGE ; STATE: PA ; COUNTRY: USA ; ZIP: 19482 ; COMPUTER READABLE FORM: MEDIUM TYPE: Diskette ; COMPUTER: IBM Compatible ; OPERATING SYSTEM: DOS ; SOFTWARE: FastSEQ for Windows Version 2.0 ; CURRENT APPLICATION DATA: APPLICATION NUMBER: US/08/986,485 ; FILING DATE: 08-DEC-1997 ; CLASSIFICATION: PRIOR APPLICATION DATA: APPLICATION NUMBER: 60/059,448 ; FILING DATE: 22-SEP-1997 ; ATTORNEY/AGENT INFORMATION: NAME: PRESTIA, PAUL F ; REGISTRATION NUMBER: 23,031 ; REFERENCE/DOCKET NUMBER: GH-70264 ; TELECOMMUNICATION INFORMATION: TELEPHONE: 610-407-0700 ; TELEFAX: 610-407-0701 ; TELEX: 846169 ; INFORMATION FOR SEQ ID NO: 2 ; SEQUENCE CHARACTERISTICS: LENGTH: 1101 amino acids ; TYPE: amino acid ; STRANDEDNESS: single ; TOPOLOGY: linear ; MOLECULE TYPE: protein ; US-08-986-485-2

Query Match 8.6%; Score 351; DB 3; Length 1101;
 Best Local Similarity 21.6%; Pred. No. 1.8e-20;
 Matches 168; Conservative 77; Mismatches 233; Indels 298; Gaps 23;

QY 30 LSESLGTLCPKSGLL-----FVPPDIDRRVTVELRGLGNFIHISRQDFANMTGLVDLT 82
 DB 193 LRSLLTLRLSKNRITQLPVRAFKL P---RLTQLDLNRRRLRLIBGLTFQGLNSLEVLUK 248
 QY 83 LSRNTISHTQPSFLDLESRLSHLSDNRLPSLGEDTLRGLVNLQHLI VNNQLGGIAD 142
 DB 249 LQNNISKLTGDAFWGLSKMHLVHLEBYDSLVEVNSGLYGLTALHQLHLSNNSIARIHRK 308
 QY 143 AFEDLLTLEDLDSYNNLHGLPWDSVRRVMVHLQSLDHLNLDHIAEGTFADLQKLARL 202
 DB 309 GW-SFCQKHELVLSPNNLRLDDESLAELSLSLRLSHNSISHIAEAGFKGLSLRVL 367
 QY 203 DLTSNRLOKLPDP--IFARSOASALTATPFAPPLS-----FSFGN--- 242
 DB 368 DLDHNEISGTIEDTSGAFSGLEPCHSKLTLFGNKIKSVAKRAFSGLEGLHNLGNATR 427
 QY 243 -----PLH-----CNCELLWL-----RRLERDDDLLETCGSPGGLKGRYF 276
 DB 428 SVQFDAFVKKMNLKELHISSDSFLCDQLKWLPPWLI GRMLQAFVVTATCAHPESLKGQSI 487
 QY 277 WHVREEFVCE-----PPLITQ----- 293
 DB 488 FSVPPSFVCDLFPQIITQPTTMMVQKDIRFTCSAASSSSPMTFAWKKNDEVLTN 547

RESULT 4
 US-09-063-950-2 ; Sequence 2, Application US/09063950C ; Patent No. 6225085 ; GENERAL INFORMATION ; APPLICANT: Holtzman, Douglas A. ; TITLE OF INVENTION: NOVEL LRSG PROTEIN AND NUCLEIC ACID MOLECULES AND USES ; FILE OF INVENTION: THEREFOR ; FILE REFERENCE: MEI-019 ; CURRENT APPLICATION NUMBER: US/09/063,950C ; CURRENT FILING DATE: 1998-04-21 ; NUMBER OF SEQ ID NOS: 9 ; SOFTWARE: Patent in Ver. 2.0 ; SEQ ID NO 2 ; LENGTH: 673 ; TYPE: PRT ; ORGANISM: Homo sapiens ; US-09-063-950-2

Query Match 7.3%; Score 289; DB 4; Length 673;
 Best Local Similarity 23.7%; Pred. No. 1.8e-16;
 Matches 169; Conservative 77; Mismatches 268; Indels 198; Gaps 28;

QY 8 LLAFGMAFVVDACPKYCVQNLSESLGTLCPKSGLLFVPPDIDRRVTVELRGLGNFIHII 67
 DB 14 LLALGPG---VQCGPSGQC---SOPQIVFCTARQTTVPRDVPDPDTVGLYFENGITML 67
 QY 68 SRQDFANMTGLVDLTLRSNTISHIQPSFLDLESRLSHLSDNRLPSLGEDTLRGLVNLQ 127
 DB 68 DAGSFAGLPGQLDLSQNIASLPSGVFQPLANLSNLDLTANRLHEITNETFRGLRRL 127
 QY 128 HLIVNNQLGGTADBAFEDFLTLLEDLDSYNNLHGLPWDSVRRVMVNLHQLSLDHNL--- 184
 DB 128 RLYLCKNRIRHIFQPAF-DTLDRLELKLQDLNELRALP---PLRPLRLLLDLSHNSLLA 183
 QY 185 -----LDHIAEGTFADLQKLARLDLTGNRLOKLPDP----- 215
 DB 184 LEPGLDITANVEALRAGLGLQDLDEGLFSRLNRLHLDVSDNQLSERVPPVIRGLRGLTR 243
 QY 216 -----PIFARSOASALTATPFAPPLSFSF-----GGNPLHCN 247
 DB 244 LRLAGNTRIAQRURPEDLAGLAALQELDVSNLSLQALPGDLSGLFPRLLRLLAARPNFNCV 303

Query Match 7.3%; Score 289; DB 4; Length 673;
 Best Local Similarity 23.7%; Pred. No. 1.8e-16;
 Matches 169; Conservative 77; Mismatches 268; Indels 198; Gaps 28;

QY 8 LLAFGMAFVVDACPKYCVQNLSESLGTLCPKSGLLFVPPDIDRRVTVELRGLGNFIHII 67
 DB 14 LLALGPG---VQCGPSGQC---SOPQIVFCTARQTTVPRDVPDPDTVGLYFENGITML 67
 QY 68 SRQDFANMTGLVDLTLRSNTISHIQPSFLDLESRLSHLSDNRLPSLGEDTLRGLVNLQ 127
 DB 68 DAGSFAGLPGQLDLSQNIASLPSGVFQPLANLSNLDLTANRLHEITNETFRGLRRL 127
 QY 128 HLIVNNQLGGTADBAFEDFLTLLEDLDSYNNLHGLPWDSVRRVMVNLHQLSLDHNL--- 184
 DB 128 RLYLCKNRIRHIFQPAF-DTLDRLELKLQDLNELRALP---PLRPLRLLLDLSHNSLLA 183
 QY 185 -----LDHIAEGTFADLQKLARLDLTGNRLOKLPDP----- 215
 DB 184 LEPGLDITANVEALRAGLGLQDLDEGLFSRLNRLHLDVSDNQLSERVPPVIRGLRGLTR 243
 QY 216 -----PIFARSOASALTATPFAPPLSFSF-----GGNPLHCN 247
 DB 244 LRLAGNTRIAQRURPEDLAGLAALQELDVSNLSLQALPGDLSGLFPRLLRLLAARPNFNCV 303

QY 248 CELLMLRERDDDDLETCGSPGKGRYFHWVREBEFVC 306
 Db 304 CPLSNFGVWVRESHV--TLASP
 QY 307 TLKCKA-----IGPPSPLIH-----WVAPDRVLGNSRSTAVYDNGTLDFI 348
 Db 346 DFGCPATITTTATVPTTRPVVREPTALSSSLATWLSL-----382
 QY 349 TTSQDSGAFCTIAANA-AGEATAMVEVSIQVPHLSNSTRTAPPKSRIS-----397
 Db 383 -----TAPATEAPSPSTAPPTVGPVQP-----QDCPSTCLNGTCHLGRH 426
 QY 398 -----DITGSSKTSRGGGSGGEPKSP-PERAVL--VSEVTTTSAIV--KWSYSK 444
 Db 427 HLAACLPEGFTGLYCESONGQTRPSPTVTRPRRSRLTGLIEPVSPTSLRVGLQRYLOG 486
 QY 445 SAPRVKMYOQYN--CSDDEVLIYRMPASNAKAFVNNILVSGTYDCLCVLAWMDDTATTL 502
 Db 487 SSVOLRSRLTYNLSGPKRLVTLRLPASLAETVTLRPNATYVCVMP-----GPCR 542
 QY 503 TATNIYVCAQFFTKADYQOCOSMH-----SOILGTMILVIGGIIVATLLVFIILMVRYKV 559
 Db 543 VPEGEACGEAHTP---PAVHSHAPVTOAREGNLPLLIAPALAAVLLAALAAVGAAYCV 599
 QY 560 CNHEAFSKMAAVSNVYQINGAQP-----PPSSAPAG--APPOG 598
 Db 600 RRGRA---MAAAADKQGVQPGAGPLEGKVPLEPGPKATGEGGEALPSG 648

RESULT 5
 US-09-191-647-2
 ; Sequence 2, Application US/09191647
 ; Patent No. 6046015
 ; GENERAL INFORMATION:
 ; APPLICANT: Goodman, Corey
 ; APPLICANT: Kid, Thomas
 ; APPLICANT: Brose, Katja
 ; APPLICANT: Tessier-Lavigne, Marc
 ; TITLE OF INVENTION: Modulating Robo: Ligand Interactions
 ; FILE REFERENCE: B98-031-3
 ; CURRENT FILING DATE: 2000-03-31
 ; PRIOR FILING DATE: 1997-11-14
 ; EARLIER APPLICATION NUMBER: 60/065,544
 ; EARLIER FILING DATE: 1997-11-14
 ; EARLIER APPLICATION NUMBER: 60/081,057
 ; EARLIER FILING DATE: 1998-04-07
 ; NUMBER OF SEQ ID NOS: 14
 ; SOFTWARE: Patent In Ver. 2.0
 ; SEQ ID NO 2
 ; LENGTH: 1525
 ; TYPE: PRT
 ; ORGANISM: human
 US-09-191-647-2

Query Match 6.9%; Score 284.5; DB 3; Length 1525;
 Best Local Similarity 29.2%; Pred. No. 1.1e-14;
 Matches 84; Conservative 36; Mismatches 117; Indels 51; Gaps 6;

QY 9 LAFGMFAFVDP-----ACPKYVCVONLSESLGTLCPKSGLLFVPPDIDRRRTVELRLGNGF 63
 Db 11 LSLGLVLAAILNKVAPQACQSCSGSTVD---CHGLALRSVPRNIPRNTERLDLNGNN 66
 QY 64 IHHISROPANMTGLVDLTLSENITSHIQPFSLDLESRLSLHLSNRLPSLGEDTLRGL 123
 Db 67 ITRITKDFAGLHRVQLMKNKISTIERGAFODKELERLNRNHLQVFPPELLFLGT 126
 QY 124 VNLQHLIVNNQLGGIADEAFEDLLEDDLSYNNLHGLPWSVRRMVLHQLSLDHN 183
 Db 127 AKLYR-----LDLSENQIOAIRKAFRGAVIDIKNLQLDYN 161
 QY 184 LLDHIAEGTFADLQKLARLDLTSNRLQKLPDPPIFARSOASALTATPFAPPL-SFSFGN 242
 Db 162 QIQCIEDGAFRALRDLEVLTLNNNITRL-----SVASFNHPKLRTRFRLHSN 209

QY 243 PLHCNCELL-----WLRRLERDDDDLETCGSPGKGRYFHWVREBEFVC 286
 Db 210 NLYCDDCHLAWLSDWLRKRPRVGLYTCMGPSHLRGNHVAEVQKREBFC 257

RESULT 6
 US-09-540-245A-2
 ; Sequence 2, Application US/09540245A
 ; Patent No. 6270984
 ; GENERAL INFORMATION:
 ; APPLICANT: Goodman, Corey
 ; APPLICANT: Kid, Thomas
 ; APPLICANT: Brose, Katja
 ; APPLICANT: Tessier-Lavigne, Marc
 ; TITLE OF INVENTION: Modulating Robo: Ligand Interactions
 ; FILE REFERENCE: B98-031-3
 ; CURRENT APPLICATION NUMBER: US/09/540,245A
 ; CURRENT FILING DATE: 2000-03-31
 ; PRIOR APPLICATION NUMBER: 60/065,544
 ; PRIOR FILING DATE: 1997-11-14
 ; PRIOR APPLICATION NUMBER: 60/081,057
 ; PRIOR FILING DATE: 1998-04-07
 ; NUMBER OF SEQ ID NOS: 20
 ; SOFTWARE: Patent In Ver. 2.0
 ; SEQ ID NO 2
 ; LENGTH: 1525
 ; TYPE: PRT
 ; ORGANISM: human
 US-09-540-245A-2

Query Match 6.9%; Score 284.5; DB 4; Length 1525;
 Best Local Similarity 29.2%; Pred. No. 1.1e-14;
 Matches 84; Conservative 36; Mismatches 117; Indels 51; Gaps 6;

QY 9 LAFGMFAFVDP-----ACPKYVCVONLSESLGTLCPKSGLLFVPPDIDRRRTVELRLGNGF 63
 Db 11 LSLGLVLAAILNKVAPQACQSCSGSTVD---CHGLALRSVPRNIPRNTERLDLNGNN 66
 QY 64 IHHISROPANMTGLVDLTLSENITSHIQPFSLDLESRLSLHLSNRLPSLGEDTLRGL 123
 Db 67 ITRITKDFAGLHRVQLMKNKISTIERGAFODKELERLNRNHLQVFPPELLFLGT 126
 QY 124 VNLQHLIVNNQLGGIADEAFEDLLEDDLSYNNLHGLPWSVRRMVLHQLSLDHN 183
 Db 127 AKLYR-----LDLSENQIOAIRKAFRGAVIDIKNLQLDYN 161
 QY 184 LLDHIAEGTFADLQKLARLDLTSNRLQKLPDPPIFARSOASALTATPFAPPL-SFSFGN 242
 Db 162 QIQCIEDGAFRALRDLEVLTLNNNITRL-----SVASFNHPKLRTRFRLHSN 209

QY 243 PLHCNCELL-----WLRRLERDDDDLETCGSPGKGRYFHWVREBEFVC 286
 Db 210 NLYCDDCHLAWLSDWLRKRPRVGLYTCMGPSHLRGNHVAEVQKREBFC 257

RESULT 7
 US-09-540-153-2
 ; Sequence 2, Application US/09540153
 ; Patent No. 6270995
 ; GENERAL INFORMATION:
 ; APPLICANT: Goodman, Corey
 ; APPLICANT: Kid, Thomas
 ; APPLICANT: Brose, Katja
 ; APPLICANT: Tessier-Lavigne, Marc
 ; TITLE OF INVENTION: Modulating Robo: Ligand Interactions
 ; FILE REFERENCE: B98-031-3
 ; CURRENT APPLICATION NUMBER: US/09/540,153
 ; CURRENT FILING DATE: 2000-03-31
 ; PRIOR APPLICATION NUMBER: 60/081,057
 ; PRIOR FILING DATE: 1998-11-13
 ; PRIOR APPLICATION NUMBER: 60/081,057
 ; PRIOR FILING DATE: 1998-04-07

```

; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 1525
; TYPE: PRT
; ORGANISM: human
US-09-540-153-2

```

```

Query Match 6.9%; Score 284.5; DB 4; Length 1525;
Best Local Similarity 29.2%; Pred. No. 1.1e-14;
Matches 84; Conservative 36; Mismatches 117; Indels 51; Gaps 6;

```

```

QY 9 LAFGMFAFVVD-----ACPKYCVQNLSLGLTLCPSKGLLFPVPPIDRRRTVELRLGGNF 63
DB 11 LSLGLVLAALNKVPAQACPAQCSGTYD-----CHGLALRSVPRNIPRNTERLDLNGNN 66
QY 64 IIHISRODFANMTGLVDTLSRNTISHIQPFSLDLESLSLHLSLHDSNRLPSLGEDTLRGL 123
DB 67 ITRIYKDFAGLRHLRVLQVMENKSTIERGAFODLKELERLRNHQLQPELLFLGT 126
QY 124 VNLQHLI VNNQLGGIADAEFEDFLTLEDLDSVNNLHGLPWSVRRVMVNLHQLSLDHN 183
DB 127 AKLYR-----LDLSENOIAIPRKA PRGAVDIKNIQLQDYN 161
QY 184 LLDHIAEGTFADLQKLARLDLTSNRLOKLPDPPIFARSOASALTATPPAPPL-SFSFGN 242
DB 162 QISCIEDGAFRALDRLEVLTLNNNITRL-----SVA SFNHMPKLRTPRLHSN 209
QY 243 PHLNCCELL-----WLRRLERDDLETCSGPGGLKGRYFWHVREBEFVC 286
DB 210 NLYCCHLAWLSDWLKRPVGLVYTCMGP SHLRGHVAVQKREFVC 257

```

```

RESULT 8
US-09-182-024A-2
; Sequence 2, Application US/09182024A
; Patent No. 6342370
; GENERAL INFORMATION:
; APPLICANT: Connolly, Timothy
; APPLICANT: Rajput, Bhanu
; TITLE OF INVENTION: Human Slit Polypeptide and Polynucleotides Encoding
; TITLE OF INVENTION: Same
; FILE REFERENCE: 640100-271
; CURRENT APPLICATION NUMBER: US/09/182,024A
; CURRENT FILING DATE: 1998-10-29
; PRIOR APPLICATION NUMBER: 60/063,946
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: 60/096,420
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 1523
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-182-024A-2

```

```

Query Match 6.5%; Score 268; DB 4; Length 1523;
Best Local Similarity 26.8%; Pred. No. 2.5e-13;
Matches 86; Conservative 41; Mismatches 130; Indels 64; Gaps 7;

```

```

QY 9 LAFGMFAFVVD-----DAPKYCVQNLSLGLTLCPSKGLLFPVPPIDRRRTVELRLGGN 62
DB 16 LALALALASVLSGPPAVACPKTKTCSAASVD-----CHGLGLRAVPRGIPRNERLDLDRN 71
QY 63 FIIHSRODFANMTGLVDTLSRNTISHIQPFSLDLESLSLHLSLHDSNRLPSLGEDTLRG 122
DB 72 NITRIKDFAGLRHLRVLQVMENKSTIERGAFODLKELERLRNHQLQPELLFLGT 131
QY 123 LVNLQHLI VNNQLGGIADAEFEDFLTLEDLDSVNNLHGLPWSVRRVMVNLHQLSLDHN 182
DB 132 TPKLRT-----LDLSENOIQIPRKA PRGAVDIKNIQLQDYN 166

```

```

QY 183 NLLDHIAGTFADLQKLARLDLTSNRLOKLPDPPIFARSOASALTATPPAPPL-SFSFGG 241
DB 167 NHISCIEDGAFRALDRLEVLTLNNNISRI-----LVT SFNHMPKIRTLRLHS 214
QY 242 NPLHCNCELL-----WLRRLERDDLETCSGPGGLKGRYFWHVREBEFVCPEPLLIQHTHK 297
DB 215 NLYCCHLAWLSDWLKRPVGLVYTCMGP SHLRGHVAVQKREFVC-----PFA 266
QY 298 LVLLEGQAATLKCKAIGDPPSP 318
DB 267 ---HSEPPSCWANSISCPS 283

```

```

RESULT 9
US-09-191-647-7
; Sequence 7, Application US/09191647
; Patent No. 6046015
; GENERAL INFORMATION:
; APPLICANT: Goodman, Corey
; APPLICANT: Kid, Thomas
; APPLICANT: Brose, Katja
; TITLE OF INVENTION: Modulating Robo: Ligand Interactions
; FILE REFERENCE: B98-031-3
; CURRENT APPLICATION NUMBER: US/09/191,647
; CURRENT FILING DATE: 1998-11-13
; EARLIER APPLICATION NUMBER: 60/065,544
; EARLIER FILING DATE: 1997-11-14
; EARLIER APPLICATION NUMBER: 60/081,057
; EARLIER FILING DATE: 1998-04-07
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 1480
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-191-647-7

```

```

Query Match 6.4%; Score 263; DB 3; Length 1480;
Best Local Similarity 28.0%; Pred. No. 6.3e-13;
Matches 89; Conservative 39; Mismatches 102; Indels 88; Gaps 11;

```

```

QY 6 GGLLAFGMFAF-----AVVDA-CPKYCVQNLSLGLTLCPSKGLLFPVPPIDRRRTVELR 58
DB 51 GGLGSVGIHPIGGGVGVITEARCPVCSCTGLNVD-----CSHRGLTSVPRKISADVERLE 106
QY 59 LGGNFIIHSRODFANMTGLVDTLSRNTISHIQPFSLDLESLSLHLSLHDSNRLPSLGED 118
DB 107 LQGNLTVIYETDFORLTKRMLQTDNQIHTIERN SFODLVSLESLERLDSNINNVITVGRR 166
QY 119 TLRGLVNLQHLI VNNQLGGIADAEFEDFLTLEDLDSVNNLHGLPWSVRRVMVNLHQL 178
DB 167 VPKGAQSLRSQLDNNQITCLDDEHAFKG-LVELEILTLNNNNLTSLP----- 212
QY 179 SLDNLLDHIAGTFADLQKLARLDLTSNRLOKLPDPPIFARSOASALTATPPAPPLSFS 238
DB 213 ---HNI-----FGLGLRALRLSDN-----PFA 233
QY 239 FGGNPLHCNCELLWLRRLERD-----DDLETCSGPGGLKGRYFWHVREBEFVCPEPLLIQHT 294
DB 234 -----CDCHLSWLSRFLRSATRLAPYTRCOSP SOLKQNVADLHLDQEFKCSG--LTEH 284
QY 295 THKLLVLEGQAATLKCKA 312
DB 285 -----APMECGA 291

```

```

RESULT 10
US-09-540-245A-7
; Sequence 7, Application US/09540245A
; Patent No. 6270984
; GENERAL INFORMATION:
; APPLICANT: Goodman, Corey

```

APPLICANT: Kid, Thomas
 APPLICANT: Brose, Katja
 APPLICANT: Tessier-Lavigne, Marc
 TITLE OF INVENTION: Modulating Robo: Ligand Interactions
 FILE REFERENCE: B98-031-3
 CURRENT APPLICATION NUMBER: US/09/540,245A
 CURRENT FILING DATE: 2000-03-31
 PRIOR APPLICATION NUMBER: 60/065,544
 PRIOR FILING DATE: 1997-11-14
 PRIOR APPLICATION NUMBER: 60/081,057
 PRIOR FILING DATE: 1998-04-07
 NUMBER OF SEQ ID NOS: 20
 SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO 7
 LENGTH: 1480
 TYPE: PRT
 ORGANISM: Drosophila melanogaster
 US-09-540-245A-7

Query Match 6.4%; Score 263; DB 4; Length 1480;
 Best Local Similarity 28.0%; Pred. No. 6.3e-13;
 Matches 89; Conservative 39; Mismatches 102; Indels 88; Gaps 11;
 QY 6 GGLLAFGMAF-----AVVDA-CPKYCVCONLSESLGTLCPKSGLLFVPPDIDRRRTVELR 58
 DB 51 GGLGSGVGIHPGGVGVITEARCPVCSCTGLNVD---CSHRGLTSVPRKISADVERLE 106
 QY 59 LGGNFIHISRODFANMTGLVDTLSRNTISHIQPFSFLDLESRLSLHLDNSRNLPSLGED 118
 DB 107 LOGNNLTVIYETDFORLTKRMLQLTNDQIHTIERNSFQDLVSLERLDSINNVIITVGRR 166
 QY 119 TRLGLVNLQHLIVNNOLGGIADAEFEDFLTLEDDLSYNNLHGLPWSVRRMVNLHQL 178
 DB 167 VFKGAQSLRSLODNNQITCLDEHAFKGL-LEVELLTLNNNLTSLP----- 212
 QY 179 SLDHNLDDHIAEGTFADLQKLARLDLTSNRLOKLPDPPFARSOASALATATPPAPPLSFS 238
 DB 213 ---HNI-----FGGLGRRLRALRUSDN-----PFA----- 233
 QY 239 FGGNPLHCNCELLMLRRLERD---DDLETGSGPGLKGRYFHVHREBEFVCEPPLITQH 294
 DB 234 -----CDCHLSWLSRFLRSATRLAPYTRCQSPSOLKQGNVADLHDQEFKCSG--LTEH 284
 QY 295 THKLLVLEGOAATLKCKA 312
 DB 285 -----APMECGA 291

RESULT 11
 US-09-540-153-7
 ; Sequence 7, Application US/09540153
 ; Patent No. 6270995
 ; GENERAL INFORMATION:
 ; APPLICANT: Goodman, Corey
 ; APPLICANT: Kid, Thomas
 ; APPLICANT: Brose, Katja
 ; APPLICANT: Tessier-Lavigne, Marc
 ; TITLE OF INVENTION: Modulating Robo: Ligand Interactions
 ; FILE REFERENCE: B98-031-3
 ; CURRENT APPLICATION NUMBER: US/09/540,153
 ; CURRENT FILING DATE: 2000-03-31
 ; PRIOR APPLICATION NUMBER: 09/191,647
 ; PRIOR FILING DATE: 1998-11-13
 ; PRIOR APPLICATION NUMBER: 60/081,057
 ; PRIOR FILING DATE: 1998-04-07
 ; NUMBER OF SEQ ID NOS: 14
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 7
 ; LENGTH: 1480
 ; TYPE: PRT
 ; ORGANISM: Drosophila melanogaster
 US-09-540-153-7

Query Match 6.4%; Score 263; DB 4; Length 1480;
 Best Local Similarity 28.0%; Pred. No. 6.3e-13;
 Matches 89; Conservative 39; Mismatches 102; Indels 88; Gaps 11;
 QY 6 GGLLAFGMAF-----AVVDA-CPKYCVCONLSESLGTLCPKSGLLFVPPDIDRRRTVELR 58
 DB 51 GGLGSGVGIHPGGVGVITEARCPVCSCTGLNVD---CSHRGLTSVPRKISADVERLE 106
 QY 59 LGGNFIHISRODFANMTGLVDTLSRNTISHIQPFSFLDLESRLSLHLDNSRNLPSLGED 118
 DB 107 LOGNNLTVIYETDFORLTKRMLQLTNDQIHTIERNSFQDLVSLERLDSINNVIITVGRR 166
 QY 119 TRLGLVNLQHLIVNNOLGGIADAEFEDFLTLEDDLSYNNLHGLPWSVRRMVNLHQL 178
 DB 167 VFKGAQSLRSLODNNQITCLDEHAFKGL-LEVELLTLNNNLTSLP----- 212
 QY 179 SLDHNLDDHIAEGTFADLQKLARLDLTSNRLOKLPDPPFARSOASALATATPPAPPLSFS 238
 DB 213 ---HNI-----FGGLGRRLRALRUSDN-----PFA----- 233
 QY 239 FGGNPLHCNCELLMLRRLERD---DDLETGSGPGLKGRYFHVHREBEFVCEPPLITQH 294
 DB 234 -----CDCHLSWLSRFLRSATRLAPYTRCQSPSOLKQGNVADLHDQEFKCSG--LTEH 284
 QY 295 THKLLVLEGOAATLKCKA 312
 DB 285 -----APMECGA 291

RESULT 12
 US-09-182-024A-5
 ; Sequence 5, Application US/09182024A
 ; Patent No. 6342370
 ; GENERAL INFORMATION:
 ; APPLICANT: Connolly, Timothy
 ; APPLICANT: Rajput, Bhanu
 ; TITLE OF INVENTION: Human Slit Polypeptide and Polynucleotides Encoding
 ; TITLE OF INVENTION: Same
 ; FILE REFERENCE: 640100-271
 ; CURRENT APPLICATION NUMBER: US/09/182,024A
 ; CURRENT FILING DATE: 1998-10-29
 ; PRIOR APPLICATION NUMBER: 60/063,946
 ; PRIOR FILING DATE: 1997-10-31
 ; PRIOR APPLICATION NUMBER: 60/096,420
 ; PRIOR FILING DATE: 1998-08-13
 ; NUMBER OF SEQ ID NOS: 5
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 5
 ; LENGTH: 1480
 ; TYPE: PRT
 ; ORGANISM: Drosophila melanogaster
 US-09-182-024A-5

Query Match 6.4%; Score 263; DB 4; Length 1480;
 Best Local Similarity 28.0%; Pred. No. 6.3e-13;
 Matches 89; Conservative 39; Mismatches 102; Indels 88; Gaps 11;
 QY 6 GGLLAFGMAF-----AVVDA-CPKYCVCONLSESLGTLCPKSGLLFVPPDIDRRRTVELR 58
 DB 51 GGLGSGVGIHPGGVGVITEARCPVCSCTGLNVD---CSHRGLTSVPRKISADVERLE 106
 QY 59 LGGNFIHISRODFANMTGLVDTLSRNTISHIQPFSFLDLESRLSLHLDNSRNLPSLGED 118
 DB 107 LOGNNLTVIYETDFORLTKRMLQLTNDQIHTIERNSFQDLVSLERLDSINNVIITVGRR 166
 QY 119 TRLGLVNLQHLIVNNOLGGIADAEFEDFLTLEDDLSYNNLHGLPWSVRRMVNLHQL 178
 DB 167 VFKGAQSLRSLODNNQITCLDEHAFKGL-LEVELLTLNNNLTSLP----- 212
 QY 179 SLDHNLDDHIAEGTFADLQKLARLDLTSNRLOKLPDPPFARSOASALATATPPAPPLSFS 238
 DB 213 ---HNI-----FGGLGRRLRALRUSDN-----PFA----- 233

QY 239 FCGNPLHCNCELLWLRLERD-----DDLTCGSPGLKGRYFHWVREBEFVCEPPLITQH 294
 Db 234 -----CDCHLSWLSRFLRSATRLAPYTRCQSPSQKQNVADLHDQEFKCG--LTEH 284
 QY 295 THKLLVLEGOAATLKCKA 312
 Db 285 -----APMECGA 291

Query Match 6.4%; Score 263; DB 5; Length 1480;
 Best Local Similarity 28.0%; Pred. No. 6.3e-13;
 Matches 89; Conservative 39; Mismatches 102; Indels 88; Gaps 11;

QY 6 GGLLAFGNMFA-----AVVDA-CPKYVCVONLSESLGTLCPFKSGLLFPVDPIDRRVVLRL 58
 Db 51 GGLGSVGIHPGGGVIVTEARCPVCSCCTGLNVD---CShRGLUTSVPRKISADVERLE 106
 QY 59 LGGNPIHISRQDFANMTGLVDLILSRNTISHIQPFPSFDLDESRLSLHLDNSRLSGLGED 118
 Db 107 LQGNLTIYETDEFQRLTKLRLMQLQTDNQIHTIERNRSFQDLVSLERLDISNNVITVQRR 166
 QY 119 TLRLGVNLQHLIVNNOGLGGIADAEFDELTLDELDSLNNLHGLPWDSVRRMVNHLQL 178
 Db 167 VFKGAQSLRSLOLDNNQITCLDERAFKG-LVELEILTUNNNLTSLP----- 212
 QY 179 SLDHNLDDHIAEGTFADLQKLRDLTNSRLOKLPDPPIFARSOASALTTATPPAPPLSFS 238
 Db 213 --HNI-----FGGLRRLRALRLSDN-----PFA----- 233
 QY 239 FCGNPLHCNCELLWLRLERD-----DDLTCGSPGLKGRYFHWVREBEFVCEPPLITQH 294
 Db 234 -----CDCHLSWLSRFLRSATRLAPYTRCQSPSQKQNVADLHDQEFKCG--LTEH 284
 QY 295 THKLLVLEGOAATLKCKA 312
 Db 285 -----APMECGA 291

RESULT 14
 PCT-US91-09055-3
 ; Sequence 3, Application PC/TUS9109055
 ; GENERAL INFORMATION:

APPLICANT: Rothberg, Jonathan Marc and Artavanis-Tsakonas, Spyridon
 TITLE OF INVENTION: Purified SLIT protein and Sequence Elements Thereof
 NUMBER OF SEQUENCES: 9
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Yale University
 ADDRESS: Office of Cooperative Research
 STREET: 246 Church Street
 CITY: New Haven
 STATE: Connecticut
 COUNTRY: USA
 ZIP: 06510
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette, 3.50 inch, 800 Kb storage
 OPERATING SYSTEM: Apple Macintosh
 SOFTWARE: Microsoft Word 4.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: PCT/US91/09055
 FILING DATE: 19911127
 CLASSIFICATION: 435
 PRIORITY APPLICATION DATA:
 APPLICATION NUMBER: US/07/624,135
 FILING DATE: 7-DEC-1990
 ATTORNEY/AGENT INFORMATION:
 NAME: Barth, Richard J.
 REGISTRATION NUMBER: 28,180
 REFERENCE/DOCKET NUMBER: 900964/RSB
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (212) 972-1400
 TELEX: 236268
 INFORMATION FOR SEQ ID NO: 3:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 222 amino acids
 TYPE: AMINO ACIDS

QY 239 FCGNPLHCNCELLWLRLERD-----DDLTCGSPGLKGRYFHWVREBEFVCEPPLITQH 294
 Db 234 -----CDCHLSWLSRFLRSATRLAPYTRCQSPSQKQNVADLHDQEFKCG--LTEH 284
 QY 295 THKLLVLEGOAATLKCKA 312
 Db 285 -----APMECGA 291

RESULT 13
 PCT-US91-09055-2
 ; Sequence 2, Application PC/TUS9109055
 ; GENERAL INFORMATION:
 APPLICANT: Rothberg, Jonathan Marc and Artavanis-Tsakonas, Spyridon
 TITLE OF INVENTION: Purified SLIT protein and Sequence Elements Thereof
 NUMBER OF SEQUENCES: 9
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Yale University
 ADDRESS: Office of Cooperative Research
 STREET: 246 Church Street
 CITY: New Haven
 STATE: Connecticut
 COUNTRY: USA
 ZIP: 06510
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette, 3.50 inch, 800 Kb storage
 OPERATING SYSTEM: Apple Macintosh
 SOFTWARE: Microsoft Word 4.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: PCT/US91/09055
 FILING DATE: 19911127
 CLASSIFICATION: 435
 PRIORITY APPLICATION DATA:
 APPLICATION NUMBER: US/07/624,135
 FILING DATE: 7-DEC-1990
 ATTORNEY/AGENT INFORMATION:
 NAME: Barth, Richard J.
 REGISTRATION NUMBER: 28,180
 REFERENCE/DOCKET NUMBER: 900964/RSB
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (212) 972-1400
 TELEX: 236268
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1480 amino acids
 TYPE: AMINO ACIDS
 TOPOLOGY: Linear
 MOLECULE TYPE: protein
 FEATURE:
 NAME/KEY: signal sequence
 LOCATION: 1 to 36
 IDENTIFICATION METHOD: similarity to other signal
 OTHER INFORMATION: Directs Export
 NAME/KEY: Four Flank-LRR-Flank domains
 LOCATION: 37 to 910
 IDENTIFICATION METHOD: Array of Flank-LRR-Flank
 OTHER INFORMATION: mediates adhesive events
 NAME/KEY: Tandem EGF-like repeats
 LOCATION: 911 to 1150
 IDENTIFICATION METHOD: similarity to tandem EGF-like
 OTHER INFORMATION: protein-protein interactions
 NAME/KEY: 7th EGF-like repeat
 LOCATION: 1353 to 1393
 IDENTIFICATION METHOD: similarity to epidermal growth
 OTHER INFORMATION: Involvement in receptor-ligand
 NAME/KEY: Alternative splice segment
 LOCATION: 1394 to 1404
 IDENTIFICATION METHOD: experimental
 OTHER INFORMATION: developmentally regulated
 NAME/KEY: COOH-terminal region

```

; TOPOLOGY: Linear
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: Flank-LRR-Flank 1
; LOCATION: 1 to 222
; IDENTIFICATION METHOD: similarity to other Flank-LRR-
; OTHER INFORMATION: mediates adhesive events
; PCT-US91-09055-3

```

```

Query Match          6.4%; Score 260.5; DB 5; Length 222;
Best Local Similarity 28.4%; Pred. No. 5.6e-14;
Matches 84; Conservative 35; Mismatches 96; Indels 81; Gaps 9;

QY 21 CPKYCVCONLSESLGTLCPKGLLFPVPPDIDRRRTVELRIGGNFIHISRQDFANWTGLVD 80
Db 1 CPRVCSCTGLNVD-----CSHRGLTSVPRKISADVERLELOGNLTVIYETDFORLTKLRM 56

QY 81 LTLRSNTIISHIQPFSELDLESRLSLHLDNRLPSLGEDTLRGLVNLQHLIVNNNOLGGIA 140
Db 57 LQITDNIQIHTIERNISFQDLVSLERLDISNNVITVGRVFKGAQSLRSLOLDNNQITCLD 116

QY 141 DEAFEDFLITLEDLDSYNNLHGLPWDSVRRMVLHQLSLDHNLLDHIAGTFFADLQKLA 200
Db 117 EHAFKG-LVELELTLNANNLTSLP-----HNI-----FGGLRLR 151

QY 201 RDLTNSRLQKLPDPFIPARSOASALTATPFAPPLSFSGNPLHCNCELLWLRRLERD- 259
Db 152 ALRLSDN-----PFA-----CDCHLSWLSRFLRSA 176

QY 260 ---DDLETCSGPGGLKGRYFHWVREEFVCEPPLITQHTHKLAVLEGQAATLKCKA 312
Db 177 TRLAPYTRCQSPQLKQNVADLHDQEFKCSG--LTEH-----APMECGA 219

```

```

RESULT 15
US-08-986-485-7
; Sequence 7, Application US/08986485
; Patent No. 6046030
; GENERAL INFORMATION:
; APPLICANT: WU, SHUJIAN
; APPLICANT: SWEET, RAYMOND
; APPLICANT: TRUNEH, ALEMSEGED
; TITLE OF INVENTION: A HUMAN LIG-1 HOMOLOG (HLIG-1)
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: RATNER & PRESTIA
; STREET: P.O. BOX 980
; CITY: VALLEY FORGE
; STATE: PA
; COUNTRY: USA
; ZIP: 19482
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/986,485
; FILING DATE: 08-DEC-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/059,448
; FILING DATE: 22-SEP-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: PRESTIA, PAUL F
; REGISTRATION NUMBER: 23,031
; REFERENCE/DOCKET NUMBER: GH-70264
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-407-0700
; TELEFAX: 610-407-0701
; TELEX: 846169
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:

```

```

; LENGTH: 231 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-986-485-7

Query Match          6.3%; Score 259; DB 3; Length 231;
Best Local Similarity 29.1%; Pred. No. 7.9e-14;
Matches 81; Conservative 33; Mismatches 94; Indels 70; Gaps 8;

QY 21 CPKYCVCONLSESLGTLCPKGLLFPVPPDIDRRRTVELRIGGNFIHISRQDFANWTGLVD 80
Db 17 CPRVCSCTGLNVD-----CSHRGLTSVPRKISADVERLELOGNLTVIYETDFORLTKLRM 72

QY 81 LTLRSNTIISHIQPFSELDLESRLSLHLDNRLPSLGEDTLRGLVNLQHLIVNNNOLGGIA 140
Db 73 LQITDNIQIHTIERNISFQDLVSLERLDISNNVITVGRVFKGAQSLRSLOLDNNQITCLD 132

QY 141 DEAFEDFLITLEDLDSYNNLHGLPWDSVRRMVLHQLSLDHNLLDHIAGTFFADLQKLA 200
Db 133 EHAFKG-LVELELTLNANNLTSLP-----HNI-----FGGLRLR 167

QY 201 RDLTNSRLQKLPDPFIPARSOASALTATPFAPPLSFSGNPLHCNCELLWLRRLERD- 259
Db 168 ALRLSDN-----PFA-----CDCHLSWLSRFLRSA 192

QY 260 ---DDLETCSGPGGLKGRYFHWVREEFVCEPPLITQHT 294
Db 193 TRLAPYTRCQSPQLKQNVADLHDQEFKCSG--LTEH 228

```

```

Search completed: June 8, 2003, 21:32:09
Job time : 41 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 8, 2003, 21:30:28 ; Search time 55 Seconds
(without alignments)
1481.029 Million cell updates/sec

Title: US-09-831-846-2
Perfect score: 4094
Sequence: 1 METLLGLLAFGMFAFVDA.....DLVGARGTGSSEWYMSTV 789

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 392085 seqs, 103240269 residues

Total number of hits satisfying chosen parameters: 392085

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications_AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/ECT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US05_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/ECTUS_PUBCOMB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match %	Length	ID	Description
1	1566	38.3	628	US-09-815-626-2	Sequence 2, Appli
2	1563.5	38.2	627	US-10-245-103-100	Sequence 100, App
3	1563.5	38.2	627	US-10-245-107-100	Sequence 100, App
4	1563.5	38.2	627	US-10-245-143-100	Sequence 100, App
5	1563.5	38.2	627	US-10-245-771-100	Sequence 100, App
6	1563.5	38.2	627	US-10-245-851-100	Sequence 100, App
7	1563.5	38.2	627	US-10-245-883-100	Sequence 100, App
8	1563.5	38.2	627	US-10-237-535-100	Sequence 100, App
9	1563.5	38.2	627	US-10-238-183-100	Sequence 100, App
10	1563.5	38.2	627	US-10-238-283-100	Sequence 100, App
11	1563.5	38.2	627	US-10-238-370-100	Sequence 100, App
12	1563.5	38.2	627	US-10-245-055-100	Sequence 100, App
13	1563.5	38.2	627	US-10-245-147-100	Sequence 100, App
14	1563.5	38.2	627	US-10-245-730-100	Sequence 100, App
15	1563.5	38.2	627	US-10-245-739-100	Sequence 100, App
16	1563.5	38.2	627	US-10-246-210-100	Sequence 100, App
17	1563.5	38.2	627	US-10-239-196-100	Sequence 100, App
18	1563.5	38.2	627	US-10-243-024-100	Sequence 100, App
19	1563.5	38.2	627	US-10-243-409-100	Sequence 100, App

20	1563.5	38.2	627	9	US-10-245-033-100	Sequence 100, App
21	1563.5	38.2	627	9	US-10-245-621-100	Sequence 100, App
22	1563.5	38.2	627	9	US-10-245-880-100	Sequence 100, App
23	1563.5	38.2	627	9	US-10-243-095-100	Sequence 100, App
24	1563.5	38.2	627	9	US-10-245-185-100	Sequence 100, App
25	1563.5	38.2	627	9	US-10-245-427-100	Sequence 100, App
26	1563.5	38.2	627	9	US-10-245-473-100	Sequence 100, App
27	1563.5	38.2	627	9	US-10-245-770-100	Sequence 100, App
28	1563.5	38.2	627	9	US-10-245-877-100	Sequence 100, App
29	1563.5	38.2	627	9	US-10-246-976-100	Sequence 100, App
30	1563.5	38.2	627	9	US-10-243-320-100	Sequence 100, App
31	1563.5	38.2	627	9	US-10-242-743-100	Sequence 100, App
32	1563.5	38.2	627	9	US-10-242-845-100	Sequence 100, App
33	1545.5	37.8	551	10	US-09-897-214-8	Sequence 8, Appli
34	1454	35.5	565	10	US-09-768-826-58	Sequence 58, Appl
35	1144	27.9	468	10	US-09-768-826-40	Sequence 40, Appl
36	452	11.0	713	9	US-10-028-392-35	Sequence 35, Appl
37	451	11.0	713	10	US-09-822-687-2	Sequence 2, Appli
38	441	10.6	634	9	US-10-028-392-2	Sequence 2, Appl
39	435.5	10.6	640	9	US-09-905-291A-292	Sequence 292, App
40	435.5	10.6	640	9	US-09-992-598-501	Sequence 501, App
41	435.5	10.6	640	9	US-09-989-293A-501	Sequence 501, App
42	435.5	10.6	640	9	US-09-902-853-292	Sequence 292, App
43	435.5	10.6	640	9	US-09-989-735-501	Sequence 501, App
44	435.5	10.6	640	9	US-09-990-444-501	Sequence 501, App
45	435.5	10.6	640	9	US-10-114-893-117	Sequence 117, App

ALIGNMENTS

RESULT 1
 US-09-815-626-2
 ; Sequence 2, Application US/09815626.
 ; Patent No. US20020076752A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Glucksmann, Maria Alexandra
 ; TITLE OF INVENTION: 33395, A-NOVEL HUMAN LEUCINE-RICH REPEAT
 ; TITLE OF INVENTION: FAMILY MEMBER AND USES-THEREOF
 ; FILE REFERENCE: 10448-031001
 ; CURRENT APPLICATION NUMBER: US/09/815,626
 ; CURRENT FILING DATE: 2001-03-23
 ; PRIOR APPLICATION NUMBER: US 60/191,863
 ; PRIOR FILING DATE: 2000-03-24
 ; NUMBER OF SEQ ID NOS: 17
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 2
 ; LENGTH: 628
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-815-626-2

Query Match 38.3%; Score 1566; DB 10; Length 628;
 Best Local Similarity 52.9%; Pred. No. 1.2e-78;
 Matches 311; Conservative 88; Mismatches 175; Indels 14; Gaps 5;
 Qy 16 AVVDACPVVCQNLSESLGLTLCPSKGLLFPVPPDIDRRTVELRLGNGFIHLSRODFAM 75
 Db 23 ATPSPCRRRCQTSLSVLCPGAGLLFPVPSLDRRAAELRLADNFASVRRDLANN 82
 Qy 76 TGLVDTLSRNTISHIQPFSELDLESRLSLHLDNSNRLPSLGEDTTLRGLVNLQHLIVNNQ 135
 Db 83 TGLHLSLSRNTIRHVAAGAFADLRALRALHLDGNRLTSLGEGQLRGLVNLRLHLSNQQ 142
 Qy 136 LGGIADAFEPPELLLELDLSYNNLHGLPWSVRMNLHOLSLDLHLLDHIAGFTPAD 195
 Db 143 LAALAAGALDDCAETLELDLSYNNLEQLPWEALGLGNVNTLGLDHLNLLASVPAGFSR 202
 Qy 196 LQKRLRLDLSNRLQKLPPIFARSOASALTATPAPPPLS---FSFGNPLHCCCELLW 252
 Db 203 LHKRLRLDMSNRLTTPDPLFSR---LPLLRPGSPASALVLAFGNPLHCCCELLW 259
 Qy 253 LRRRLRDLLETGCGFGLKGRYFWHVREEFVCEPPLJITQHTHKLLELQGAATLKCKA 312

```

Db 260 LRLAREDDLEACAGPPALGGGRYFWAGVEEVEFCVPPVTHRSPLAVPAGRPAALRCRA 319
Qy 313 IGDPSPLIHWVADPDLVGNSSRTAVYDNGTLDIFITTSQDSGAFTCIAANAAGEATAMV 372
Db 320 VGDPEPRVRVWVSGPQRLGNSSRARAFNGTLELLVTEPGDGGIFTCIAANAAGEATAAV 379
Qy 373 EVSI--VOLPHLSNSTRTAPPKSRSLDITGSSKTSRGGGSGGGEPPKSPPERAVLVE 430
Db 380 ELTVGPPPPQLANSTSCDPPRDGPDALTPPSAASASAKVADTG----PPTDRGVQVTE 435
Qy 431 VTTTALSALVWVSKSAPRVKMYQLOVNCSDDEVLIVYRMIIPASNKAFVNNVILVSGTYDLC 490
Db 436 HGATAALVQWDPQRPFGIRMYQIYNSSADDILVYRMIIPASRSFLLTDLASGRTYDLC 495
Qy 491 VLAMWDDTATTLTATNIVGCAOFFTKADYPOCOSMHSQILGGTMIIVGIIIVATLLVFI 550
Db 496 VLAVVGSATGLTATRPVGCARFSTEPALRPCGAPHAPFLGTTMIILGGVIVASVLVFI 555
Qy 551 VILMRYKVCNHEAP--SKMAAAVSNVYQTNCAQPPPPSSAPAGAPP 596
Db 556 FVLLMRYKVGQPPGKAKIPAPVSSVCSQTNALGPTPTAPPAPPEP 603

```

```

RESULT 2
US-10-245-103-100
; Sequence 100, Application US/10245103
; Publication No. US2003068778A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin
; APPLICANT: Eaton, Dan
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Goddard, Audrey
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Phillippe
; APPLICANT: Watanabe, Colin
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; APPLICANT: Fong, Sherman
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3630R1C112
; CURRENT APPLICATION NUMBER: US/10/245,103
; CURRENT FILING DATE: 2002-09-17
; PRIOR APPLICATION NUMBER: 10/197942
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/059114
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/063046
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/065027
; PRIOR FILING DATE: 1997-11-10
; PRIOR APPLICATION NUMBER: 60/079689
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/086478
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090689
; PRIOR FILING DATE: 1998-06-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 116
; SEQ ID NO 100
; LENGTH: 627
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-245-103-100

```

```

Query Match 38.2%; Score 1563.5; DB 9; Length 627;
Best Local Similarity 53.1%; Pred. No. 1.7e-78;
Matches 312; Conservative 88; Mismatches 173; Indels 15; Gaps 6;
Qy 16 AVVDAPKYCVCONLESIGTLCPKSGLLLPVPPDIDRRIVELRGNFIHISRODFANM 75
Db 23 ATPSPCRRRCRQTSQPLSVLFCGAGLFFVPSLDRRAAELRLADNFIVASRRRDLJAN 82
Qy 76 TGLVDLTLSENTHSHIOPFPFLDLESRLSHLDNSRSLPGLGEDTLRGLVNLQHLIVNNO 135
Db 83 TGLHLHLSRNTIRHVAAGAFADRLRALHLDGNRLTSLGEGQLRGLVNLHLLSNNQ 142
Qy 136 LGGIADFAFDLITLEDLDSYNNLHGLPMDVSRVMVNLHQLSLDHNLDHIAEGTFAD 195
Db 143 LAALAGALDDCAETLEDLDSYNNLHGLPMDVSRVMVNLHQLSLDHNLDHIAEGTFAD 201
Qy 196 LQKLARLDLTSNRLOKLPDPIPARSOASALTATFPAPPLS---FSGGNPLHCNCELLW 252
Db 202 LHKLARLDLTSNRLLTTPDPLFSR---LPELLARPRGSPASALVAFGGNPLHCNCELLW 258
Qy 253 LRLERDDLETGSGPGLKGRYFMHVRREEFVCEPEPLITQHTHKLVLLEGQAATLKCKA 312
Db 259 LRLAREDDLEACASPPALGGGRYFWAGVEEVEFCVPPVTHRSPLAVPAGRPAALRCRA 318
Qy 313 IGDPSPLIHWVADPDLVGNSSRTAVYDNGTLDIFITTSQDSGAFTCIAANAAGEATAMV 372
Db 319 VGDPEPRVRVWVSGPQRLGNSSRARAFNGTLELLVTEPGDGGIFTCIAANAAGEATAAV 378
Qy 373 EVSI--VOLPHLSNSTRTAPPKSRSLDITGSSKTSRGGGSGGGEPPKSPPERAVLVE 430
Db 379 ELTVGPPPPQLANSTSCDPPRDGPDALTPPSAASASAKVADTG----PPTDRGVQVTE 434
Qy 431 VTTTALSALVWVSKSAPRVKMYQLOVNCSDDEVLIVYRMIIPASNKAFVNNVILVSGTYDLC 490
Db 435 HGATAALVQWDPQRPFGIRMYQIYNSSADDILVYRMIIPASRSFLLTDLASGRTYDLC 494
Qy 491 VLAMWDDTATTLTATNIVGCAOFFTKADYPOCOSMHSQILGGTMIIVGIIIVATLLVFI 550
Db 495 VLAVVGSATGLTATRPVGCARFSTEPALRPCGAPHAPFLGTTMIILGGVIVASVLVFI 554
Qy 551 VILMRYKVCNHEAP--SKMAAAVSNVYQTNCAQPPPPSSAPAGAPP 596
Db 556 FVLLMRYKVGQPPGKAKIPAPVSSVCSQTNALGPTPTAPPAPPEP 602

```

```

RESULT 3
US-10-245-107-100
; Sequence 100, Application US/10245107
; Publication No. US2003068779A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin
; APPLICANT: Eaton, Dan
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Goddard, Audrey
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Phillippe
; APPLICANT: Watanabe, Colin
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; APPLICANT: Fong, Sherman
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3630R1C71
; CURRENT APPLICATION NUMBER: US/10/245,107
; CURRENT FILING DATE: 2002-09-16
; PRIOR APPLICATION NUMBER: 10/197942
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/059114
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/063046

```



```

; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/065027
; PRIOR FILING DATE: 1997-11-10
; PRIOR APPLICATION NUMBER: 60/079689
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/086478
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090689
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 116
; SEQ ID NO 100
; LENGTH: 627
; TYPE: PRT
; ORGANISM: Homo Sapien
; US-10-245-107-100

Query Match      38.2%; Score 1563.5; DB 9; Length 627;
Best Local Similarity 53.1%; Pred. No. 1.7e-78;
Matches 312; Conservative 88; Mismatches 173; Indels 15; Gaps 6;

QY 16 AVVDACPKYVCQNLSSEGLTLCPSKGLLFPVPPIDRRVLRLLGGNFIHISRODFANM 75
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
23 ATPSPCRRRCQTSQSLPSLPCGAGLLFVPPSLDRRAAELRLADNFIAFVRRDRLANM 82
QY 76 TGLVDLTLRNTISHIQPFSLDLESRLSHLDSNRLPSLGEDTLRGLVNLQHLVNNQ 135
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
83 TGLLHLSLRNTIRHVAAGAFADRALRALHLDGNRLTSLGEGQLRGLVNLRHLLSNNQ 142
QY 136 LGGIADAEFEDLLEDLDLSDYNNLHGLPWDSVRRMNLHOLSJDHLLDHIAGTFAD 195
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
143 LAAALAGALDDCAETLEDLSDYNNLEQLPWEALGRGNVNTGLDHNLLASV-PCAFSR 201
QY 196 LOKLARLDLTSNRLOKLPDPFARSAQSALTATFPAPPLS---FSFGNPLHNCCELLW 252
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
202 LHKLARLDWTSNRLTTPDPPLFSR--LPLLARPGSPASALVAFGPNLHCCELVW 258
QY 253 LRLRERDDLETGSPGGLKGRYFHWVREEEFCVPPPLITQHTKLLVLEGOAATLKCA 312
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
259 LRLAREDDLEACASPPALGGRYFVAWGEEFVCEPPVTVHRSPLAVPAGRPAAALRCRA 318
QY 313 IGDPSPLHWVAPDRLVGNSSRTAVYDNGTLDIFITTSQDSGATFCIAANAAGETAAMV 372
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
319 VGDPEPRVWRVSPQGRLLGNSSRAAFPNGLTLELLVTFEPDGGIFTCIAANAAGETAAMV 378
QY 373 EYSI--VQLPHLSNSTSRTPPKRSRLSDITGSSKTSRGGGSGGGEPKSPPERAVLSE 430
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
379 ELTVGPPPPQLANSTSCDPRDGDPPDALTPSAASAKVADTG---PPTRDGRVQVTE 434
QY 431 VTTTSALVKWSKSAAPRYKMYQLQVNCSDDEVLIVYRMIYPAKNAFVNNVLSGTYDLC 490
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
435 HGATAALVQWPPQPIPIRMIYQVYNSADDILVYRMIYPAESRSFLLDLASGRTYDLC 494
QY 491 VLAMWDDTATLTAIIVCCAFFYKADYPCQOSMHSQILGGTMILVIGGIIVATLLVPI 550
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
495 VLAVYEDSATGTATRPPVGCARFTEPALRPCAPHAPLGGTMIALGSGVIIVASVYVFI 554
QY 551 VTLVMRYKVCNHEAP--SKWAAAVSNVYSQTNGAQPPPPSSSAGAPAGAPP 596
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
555 FVLLMRYKHGGQPFGKAKIPAVSVCSQTNGALGTPTPAPPAPPEP 602

RESULT 4
US-10-245-143-100
; Sequence 100, Application US/10245143
; Publication No. US20030068780A1
; GENERAL INFORMATION:

```

```

; APPLICANT: Baker, Kevin
; APPLICANT: Eaton, Dan
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Goddard, Audrey
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Phillippe
; APPLICANT: Watanabe, Collin
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; APPLICANT: Fong, Sherman
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3630RIC90
; CURRENT APPLICATION NUMBER: US/10/245.143
; CURRENT FILING DATE: 2002-09-16
; PRIOR APPLICATION NUMBER: 10/197942
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/059114
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/063046
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/065027
; PRIOR FILING DATE: 1997-11-10
; PRIOR APPLICATION NUMBER: 60/079689
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/086478
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090689
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 116
; SEQ ID NO 100
; LENGTH: 627
; TYPE: PRT
; ORGANISM: Homo Sapien
; US-10-245-143-100

Query Match      38.2%; Score 1563.5; DB 9; Length 627;
Best Local Similarity 53.1%; Pred. No. 1.7e-78;
Matches 312; Conservative 88; Mismatches 173; Indels 15; Gaps 6;

QY 16 AVVDACPKYVCQNLSSEGLTLCPSKGLLFPVPPIDRRVLRLLGGNFIHISRODFANM 75
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
23 ATPSPCRRRCQTSQSLPSLPCGAGLLFVPPSLDRRAAELRLADNFIAFVRRDRLANM 82
QY 76 TGLVDLTLRNTISHIQPFSLDLESRLSHLDSNRLPSLGEDTLRGLVNLQHLVNNQ 135
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
83 TGLLHLSLRNTIRHVAAGAFADRALRALHLDGNRLTSLGEGQLRGLVNLRHLLSNNQ 142
QY 136 LGGIADAEFEDLLEDLDLSDYNNLHGLPWDSVRRMNLHOLSJDHLLDHIAGTFAD 195
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
143 LAAALAGALDDCAETLEDLSDYNNLEQLPWEALGRGNVNTGLDHNLLASV-PCAFSR 201
QY 196 LOKLARLDLTSNRLOKLPDPFARSAQSALTATFPAPPLS---FSFGNPLHNCCELLW 252
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
202 LHKLARLDWTSNRLTTPDPPLFSR--LPLLARPGSPASALVAFGPNLHCCELVW 258
QY 253 LRLRERDDLETGSPGGLKGRYFHWVREEEFCVPPPLITQHTKLLVLEGOAATLKCA 312
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
259 LRLAREDDLEACASPPALGGRYFVAWGEEFVCEPPVTVHRSPLAVPAGRPAAALRCRA 318
QY 313 IGDPSPLHWVAPDRLVGNSSRTAVYDNGTLDIFITTSQDSGATFCIAANAAGETAAMV 372
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
319 VGDPEPRVWRVSPQGRLLGNSSRAAFPNGLTLELLVTFEPDGGIFTCIAANAAGETAAMV 378

```

QY 373 EVSI--VOLPHLSNSTSRTAPPKRSLSLDTGSSKTSRGGGGGGGGPPKSPPERAVLSE 430
 Db 379 ELTVGPPPPOLANSTSCDPPRDGDDPALTTPPSAASAKAVDTG---PPTDRGVQVTE 434
 QY 431 VTTTSALVKWSKSAAPRVKMYQYNCSDDVLIYRMIPASNKAFVNNLVSGTYDLC 490
 Db 435 HGATAALVQWPORPIPGIRMYQIYNSSADLLVYRMIPAESRSFLITDASGRTYDLC 494
 QY 491 VLAMWDDTATLTATNIVGCAOFFTKADYPOCOSMHSQILGGTMIIVIGIIVATLLVFI 550
 Db 495 VLAVYEDSATGLTATRPVGCARFSTEPALRPCGAPHFLGGTMIIVIGIIVASVLVFI 554
 QY 551 VILVMRYKVCNHEAP--SKMAAAVSNVYVYOTNGAOPPSSAPAGAPP 596
 Db 555 FVLLMRYKHGGQPPGKAKIPAPVSSVCSQTNALGPTPTAPPAPER 602

RESULT 5
 US-10-245-771-100
 ; Sequence 100, Application US/10245771
 ; Publication No. US20030068781A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Baker, Kevin
 ; APPLICANT: Eaton, Dan
 ; APPLICANT: Filvaroff, Ellen
 ; APPLICANT: Goddard, Audrey
 ; APPLICANT: Grimaldi, J. Christopher
 ; APPLICANT: Gurney, Austin
 ; APPLICANT: Smith, Victoria
 ; APPLICANT: Stephan, Jean-Phillippe
 ; APPLICANT: Watanabe, Colin
 ; APPLICANT: Wood, William
 ; APPLICANT: Zhang, Zemin
 ; APPLICANT: Fong, Sherman
 ; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
 ; TITLE OF INVENTION: ACIDS ENCODING THE SAME
 ; FILE REFERENCE: P3630R1C98
 ; CURRENT APPLICATION NUMBER: US/10/245,771
 ; CURRENT FILING DATE: 2002-09-16
 ; PRIOR APPLICATION NUMBER: 10/197942
 ; PRIOR FILING DATE: 2002-07-18
 ; PRIOR APPLICATION NUMBER: 60/059114
 ; PRIOR FILING DATE: 1997-09-17
 ; PRIOR APPLICATION NUMBER: 60/063046
 ; PRIOR FILING DATE: 1997-10-24
 ; PRIOR APPLICATION NUMBER: 60/065027
 ; PRIOR FILING DATE: 1997-11-10
 ; PRIOR APPLICATION NUMBER: 60/079689
 ; PRIOR FILING DATE: 1998-03-27
 ; PRIOR APPLICATION NUMBER: 60/086478
 ; PRIOR FILING DATE: 1998-05-22
 ; PRIOR APPLICATION NUMBER: 60/087607
 ; PRIOR FILING DATE: 1998-06-02
 ; PRIOR APPLICATION NUMBER: 60/089801
 ; PRIOR FILING DATE: 1998-06-18
 ; PRIOR APPLICATION NUMBER: 60/090557
 ; PRIOR FILING DATE: 1998-06-24
 ; PRIOR APPLICATION NUMBER: 60/090689
 ; PRIOR FILING DATE: 1998-06-25
 ; Remaining prior Application data removed - See File Wrapper or PALM.
 ; NUMBER OF SEQ ID NOS: 116
 ; LENGTH: 627
 ; TYPE: PRT
 ; ORGANISM: Homo Sapien
 ; US-10-245-771-100

Query Match 38.2%; Score 1563.5; DB 9; Length 627;
 Best Local Similarity 53.1%; Pred. No. 1.7e-78;
 Matches 312; Conservative 88; Mismatches 173; Indels 15; Gaps 6;
 QY 16 AVVDAPKVCVQNLSGLTICPSKGLLFPVPPIDRRVVELRNLGNFIIHISRQDFANM 75

Db 23 APPSPCRRRCQTCQSLFLSLVCPGAGLLFVPPSLDRRAAERLADNFIAVRRRDLANM 82
 QY 76 TGLVDLTSRNTISHIQPPFLDLSRSLHLSNRFLSGEDLTLRGLVNLQHLIWNNOQ 135
 Db 83 TGLLHLSRNTIRHVAAGAFADLRALRHLGDLGNRLTSLGEGQLRGLVNLRHLLISNOQ 142
 QY 136 LGGIADAFEDFLLELDLDSYNNLHGLPWDSVRMVNHLQSLDHNLLDHIAGGTAD 195
 Db 143 LAALAAGALDCAETLELDLDSYNNLEQLPWEALGRGLGNVNTLGDHNLASV-PCAFSR 201
 QY 196 LQKLARLDLTSNRLOKLPDPFARSQASALTATPAPLPS--FSFGGNPLHNCCELLW 252
 Db 202 LHKLARLDMTNSRLTTIPDPLFSR---LPLLARPRGSPASALVLAFCGNPLHNCCELLW 258
 QY 253 LRRLERDDDLTSCGSPGLKGRYFHWVREBFFVCEPPLITQIHTHKLKLLVLEGOAATLKCKA 312
 Db 259 LRLAREDDLEACASPPALGGRYFWAVGEEFVCEPVPVTHRSPLAVPAGRPAALRCRA 318
 QY 313 IGDPSPLIHWAPDDRLVGNSSRTAVYDNGTLDIFITTSQDSGAFTCIAANAGATAMV 372
 Db 319 VGDPEPRVWVSPQGRLLGNSSRARAFPNGTLELVTEPPDGGGIPTCIAANAAAGATAAV 378
 QY 373 EVSI--VOLPHLSNSTSRTAPPKRSLSLDTGSSKTSRGGGGGGGGPPKSPPERAVLSE 430
 Db 379 ELTVGPPPPOLANSTSCDPPRDGDDPALTTPPSAASAKAVDTG---PPTDRGVQVTE 434
 QY 431 VTTTSALVKWSKSAAPRVKMYQYNCSDDVLIYRMIPASNKAFVNNLVSGTYDLC 490
 Db 435 HGATAALVQWPORPIPGIRMYQIYNSSADLLVYRMIPAESRSFLITDASGRTYDLC 494
 QY 491 VLAMWDDTATLTATNIVGCAOFFTKADYPOCOSMHSQILGGTMIIVIGIIVATLLVFI 550
 Db 495 VLAVYEDSATGLTATRPVGCARFSTEPALRPCGAPHFLGGTMIIVIGIIVASVLVFI 554
 QY 551 VILVMRYKVCNHEAP--SKMAAAVSNVYVYOTNGAOPPSSAPAGAPP 596
 Db 555 FVLLMRYKHGGQPPGKAKIPAPVSSVCSQTNALGPTPTAPPAPER 602

RESULT 6
 US-10-245-851-100
 ; Sequence 100, Application US/10245851
 ; Publication No. US20030068782A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Baker, Kevin
 ; APPLICANT: Eaton, Dan
 ; APPLICANT: Filvaroff, Ellen
 ; APPLICANT: Goddard, Audrey
 ; APPLICANT: Grimaldi, J. Christopher
 ; APPLICANT: Gurney, Austin
 ; APPLICANT: Smith, Victoria
 ; APPLICANT: Stephan, Jean-Phillippe
 ; APPLICANT: Watanabe, Colin
 ; APPLICANT: Wood, William
 ; APPLICANT: Zhang, Zemin
 ; APPLICANT: Fong, Sherman
 ; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
 ; TITLE OF INVENTION: ACIDS ENCODING THE SAME
 ; FILE REFERENCE: P3630R1C93
 ; CURRENT APPLICATION NUMBER: US/10/245,851
 ; CURRENT FILING DATE: 2002-09-16
 ; PRIOR APPLICATION NUMBER: 10/197942
 ; PRIOR FILING DATE: 2002-07-18
 ; PRIOR APPLICATION NUMBER: 60/059114
 ; PRIOR FILING DATE: 1997-09-17
 ; PRIOR APPLICATION NUMBER: 60/063046
 ; PRIOR FILING DATE: 1997-10-24
 ; PRIOR APPLICATION NUMBER: 60/065027
 ; PRIOR FILING DATE: 1997-11-10
 ; PRIOR APPLICATION NUMBER: 60/079689
 ; PRIOR FILING DATE: 1998-03-27
 ; PRIOR APPLICATION NUMBER: 60/086478
 ; PRIOR FILING DATE: 1998-05-22

; PRIOR APPLICATION NUMBER: 60/087607
 ; PRIOR FILING DATE: 1998-06-02
 ; PRIOR APPLICATION NUMBER: 60/089801
 ; PRIOR FILING DATE: 1998-06-18
 ; PRIOR APPLICATION NUMBER: 60/090557
 ; PRIOR FILING DATE: 1998-06-24
 ; PRIOR APPLICATION NUMBER: 60/090689
 ; PRIOR FILING DATE: 1998-06-25
 ; Remaining Prior Application data removed - See File Wrapper or PALM.
 ; NUMBER OF SEQ ID NOS: 116
 ; SEQ ID NO 100
 ; LENGTH: 627
 ; TYPE: PRT
 ; ORGANISM: Homo Sapien
 ; US-10-245-851-100

Query Match 38.2%; Score 1563.5; DB 9; Length 627;
 Best Local Similarity 53.1%; Pred. No. 1.7e-78;
 Matches 312; Conservative 88; Mismatches 173; Indels 15; Gaps 6;

QY	16	AVVDACPKYVCQNLSESLGTLCPKSGLLFVPPDIDRRVLELRGLGNFIHISRODFANM	75
Db	23	ATPSPCPRRCRCOTQSLPLSLVCPGAGLLFVPPSLDRRAAELRLADNFIAVRRRDLANM	82
QY	76	TGLVDLTLSRNTISHIQPFSLDLSRLHLDNSRLSLGDETLRGLVNLQHLIVNNQ	135
Db	83	TGLLHLSLRNTRIRVAAGAFADLRALRALHLDGNRLTSLGEGQLRGLVNLRLHLLSNQ	142
QY	136	LGGIADAEFEDLLELDLSDYNNLHGLPWDSVRRMNLHQLSLDHLNLLDHIAGTFAD	195
Db	143	LAALAAGALDDCAETLEDLSDYNNLEQLPWEALGRGLGNVNTLGLDHNLLASV-PGAFSR	201
QY	196	LQKLARLDLTSNRLOKLPDPPIFARSQASALATPAPPPLS---FSFGNPLHCNCELLW	252
Db	202	LHKLARLDWTSNRLTTPDPPLFSR--LPLLARPRGSPASALVLAFCGNPLHCNCELLW	258
QY	253	LRLRLRDDLETCGSPGKGRYFHWVREBEFVCEPPLITQTHKLLVLEGOAATLKCKA	312
Db	259	LRLRLRDDLEACASPPALGGRRYFVAWVGEFVCEPVPVTHRSPPPLAVPAGRPALRCRA	318
QY	313	IGDPSPLIHWAPDDRLVGNSSRTAVYDNGTLDIFITTSQDSGAFTCIAANAAGEATMV	372
Db	319	VGDPEPRVWVSPQGLLGNSSRARAFPNGLTELLVTEPCDGGIFTCIAANAAGEATAAV	378
QY	373	EVSI--VOLPHLSNSTSRAPPKSLSDITGSKTSRGGGGGGGPPKSPPERAVLVE	430
Db	379	ELTVGPPPPQLANSTSCDPRDGDALTPPSAASAKVADTG---PPTDRGVQVTE	434
QY	431	VTTTSALVKWSVKSAPRVKMYQLOYNCSDDDEVLYIRMI PASNKAFVNNLVSGTYDLC	490
Db	435	HGATAALVQMPDQRPPIGIRMYQIQVNSSADDILVYRMI PAESRSPLLTDLASGRTYDLC	494
QY	491	VLAMWDDTATLTATNIVCAQFFTKADYPCQSMHSOILGGTMILVIGIIVATLLVFI	550
Db	495	VLAIVEDSATGATRPVCGARFSTEPALRPCGAPHAPFLGGTMIHALGGVIVASVLVFI	554
QY	551	VILMRYKVNHEAP--SKMAAAVSNVYQTNAGAAPPPSSAPAGAPP	596
Db	555	FVLLMRYKHVGGPPCKAKIPAPVSSVCQTNGALGPTTPAPPAPPEP	602

RESULT 7

US-10-245-883-100
 ; Sequence 100, Application US/10245883
 ; Publication No. US20030068783A1
 ; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin
 ; APPLICANT: Eaton, Dan
 ; APPLICANT: Filvaroff, Ellen
 ; APPLICANT: Goddard, Audrey
 ; APPLICANT: Grimaldi, J. Christopher
 ; APPLICANT: Gurney, Austin
 ; APPLICANT: Smith, Victoria

; APPLICANT: Stephan, Jean-Phillippe
 ; APPLICANT: Watanbe, Colin
 ; APPLICANT: Wood, William
 ; APPLICANT: Zhang, Zemin
 ; APPLICANT: Fong, Sherman
 ; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
 ; TITLE OF INVENTION: ACIDS ENCODING THE SAME
 ; FILE REFERENCE: P3630RIC70
 ; CURRENT APPLICATION NUMBER: US/10/245,883
 ; CURRENT FILING DATE: 2002-09-16
 ; PRIOR APPLICATION NUMBER: 10/197942
 ; PRIOR FILING DATE: 2002-07-18
 ; PRIOR APPLICATION NUMBER: 60/059114
 ; PRIOR FILING DATE: 1997-09-17
 ; PRIOR APPLICATION NUMBER: 60/063046
 ; PRIOR FILING DATE: 1997-10-24
 ; PRIOR APPLICATION NUMBER: 60/065027
 ; PRIOR FILING DATE: 1997-11-10
 ; PRIOR APPLICATION NUMBER: 60/079689
 ; PRIOR FILING DATE: 1998-03-27
 ; PRIOR APPLICATION NUMBER: 60/086478
 ; PRIOR FILING DATE: 1998-05-22
 ; PRIOR APPLICATION NUMBER: 60/087607
 ; PRIOR FILING DATE: 1998-06-02
 ; PRIOR APPLICATION NUMBER: 60/089801
 ; PRIOR FILING DATE: 1998-06-18
 ; PRIOR APPLICATION NUMBER: 60/090557
 ; PRIOR FILING DATE: 1998-06-24
 ; PRIOR APPLICATION NUMBER: 60/090689
 ; PRIOR FILING DATE: 1998-06-25
 ; Remaining Prior Application data removed - See File Wrapper or PALM.
 ; NUMBER OF SEQ ID NOS: 116
 ; SEQ ID NO 100
 ; LENGTH: 627
 ; TYPE: PRT
 ; ORGANISM: Homo Sapien
 ; US-10-245-883-100

Query Match 38.2%; Score 1563.5; DB 9; Length 627;
 Best Local Similarity 53.1%; Pred. No. 1.7e-78;
 Matches 312; Conservative 88; Mismatches 173; Indels 15; Gaps 6;

QY	16	AVVDACPKYVCQNLSESLGTLCPKSGLLFVPPDIDRRVLELRGLGNFIHISRODFANM	75
Db	23	ATPSPCPRRCRCOTQSLPLSLVCPGAGLLFVPPSLDRRAAELRLADNFIAVRRRDLANM	82
QY	76	TGLVDLTLSRNTISHIQPFSLDLSRLHLDNSRLSLGDETLRGLVNLQHLIVNNQ	135
Db	83	TGLLHLSLRNTRIRVAAGAFADLRALRALHLDGNRLTSLGEGQLRGLVNLRLHLLSNQ	142
QY	136	LGGIADAEFEDLLELDLSDYNNLHGLPWDSVRRMNLHQLSLDHLNLLDHIAGTFAD	195
Db	143	LAALAAGALDDCAETLEDLSDYNNLEQLPWEALGRGLGNVNTLGLDHNLLASV-PGAFSR	201
QY	196	LQKLARLDLTSNRLOKLPDPPIFARSQASALATPAPPPLS---FSFGNPLHCNCELLW	252
Db	202	LHKLARLDWTSNRLTTPDPPLFSR--LPLLARPRGSPASALVLAFCGNPLHCNCELLW	258
QY	253	LRLRLRDDLETCGSPGKGRYFHWVREBEFVCEPPLITQTHKLLVLEGOAATLKCKA	312
Db	259	LRLRLRDDLEACASPPALGGRRYFVAWVGEFVCEPVPVTHRSPPPLAVPAGRPALRCRA	318
QY	313	IGDPSPLIHWAPDDRLVGNSSRTAVYDNGTLDIFITTSQDSGAFTCIAANAAGEATMV	372
Db	319	VGDPEPRVWVSPQGLLGNSSRARAFPNGLTELLVTEPCDGGIFTCIAANAAGEATAAV	378
QY	373	EVSI--VOLPHLSNSTSRAPPKSLSDITGSKTSRGGGGGGGPPKSPPERAVLVE	430
Db	379	ELTVGPPPPQLANSTSCDPRDGDALTPPSAASAKVADTG---PPTDRGVQVTE	434
QY	431	VTTTSALVKWSVKSAPRVKMYQLOYNCSDDDEVLYIRMI PASNKAFVNNLVSGTYDLC	490
Db	435	HGATAALVQMPDQRPPIGIRMYQIQVNSSADDILVYRMI PAESRSPLLTDLASGRTYDLC	494

; PRIOR APPLICATION NUMBER: 60/235147
 ; PRIOR FILING DATE: 2000-09-22
 ; PRIOR APPLICATION NUMBER: 60/261878
 ; PRIOR FILING DATE: 2001-01-12
 ; PRIOR APPLICATION NUMBER: 60/261910
 ; PRIOR FILING DATE: 2001-01-16
 ; PRIOR APPLICATION NUMBER: 60/261939
 ; PRIOR FILING DATE: 2001-01-16
 ; PRIOR APPLICATION NUMBER: 60/262150
 ; PRIOR FILING DATE: 2001-01-16
 ; PRIOR APPLICATION NUMBER: 60/264395
 ; PRIOR FILING DATE: 2001-01-25
 ; PRIOR APPLICATION NUMBER: 60/266421
 ; PRIOR FILING DATE: 2001-02-02
 ; PRIOR APPLICATION NUMBER: 60/267623
 ; PRIOR FILING DATE: 2001-02-09
 ; PRIOR APPLICATION NUMBER: 60/274399
 ; PRIOR FILING DATE: 2001-03-09
 ; PRIOR APPLICATION NUMBER: 60/280982
 ; PRIOR FILING DATE: 2001-04-03
 ; PRIOR APPLICATION NUMBER: 60/282129
 ; PRIOR FILING DATE: 2001-04-04
 ; PRIOR APPLICATION NUMBER: 60/282199
 ; PRIOR FILING DATE: 2001-04-04
 ; PRIOR APPLICATION NUMBER: 60/290589
 ; PRIOR FILING DATE: 2001-05-09
 ; PRIOR APPLICATION NUMBER: 09/180997
 ; PRIOR FILING DATE: 1998-11-19
 ; PRIOR APPLICATION NUMBER: 09/267213
 ; PRIOR FILING DATE: 1999-03-12
 ; PRIOR APPLICATION NUMBER: 09/380137
 ; PRIOR FILING DATE: 1999-08-25
 ; PRIOR APPLICATION NUMBER: 09/380138
 ; PRIOR FILING DATE: 1999-08-25
 ; PRIOR APPLICATION NUMBER: 09/403297
 ; PRIOR FILING DATE: 1999-10-16
 ; PRIOR APPLICATION NUMBER: 09/423741
 ; PRIOR FILING DATE: 1999-11-10
 ; PRIOR APPLICATION NUMBER: 09/709238
 ; PRIOR FILING DATE: 2000-11-08
 ; PRIOR APPLICATION NUMBER: 09/802706
 ; PRIOR FILING DATE: 2001-03-09
 ; PRIOR APPLICATION NUMBER: 09/872035
 ; PRIOR FILING DATE: 2001-06-01
 ; PRIOR APPLICATION NUMBER: 09/918585
 ; PRIOR FILING DATE: 2001-07-30
 ; PRIOR APPLICATION NUMBER: 09/924419
 ; PRIOR FILING DATE: 2001-08-06
 ; PRIOR APPLICATION NUMBER: 09/927796
 ; PRIOR FILING DATE: 2001-08-09
 ; PRIOR APPLICATION NUMBER: 09/929404
 ; PRIOR FILING DATE: 2001-08-13
 ; PRIOR APPLICATION NUMBER: 09/931836
 ; PRIOR FILING DATE: 2001-08-16
 ; PRIOR APPLICATION NUMBER: 09/941992
 ; PRIOR FILING DATE: 2001-08-28
 ; PRIOR APPLICATION NUMBER: 09/946374
 ; PRIOR FILING DATE: 2001-09-04
 ; PRIOR APPLICATION NUMBER: 10/001054
 ; PRIOR FILING DATE: 2001-11-30
 ; PRIOR APPLICATION NUMBER: 10/052586
 ; PRIOR FILING DATE: 2002-01-15
 ; PRIOR APPLICATION NUMBER: 10/081056
 ; PRIOR FILING DATE: 2002-02-20
 ; PRIOR APPLICATION NUMBER: 10/119480
 ; PRIOR FILING DATE: 2002-04-09

Query Match 38.2%; Score 1563.5; DB 9; Length 627;
 Best Local Similarity 53.1%; Pred. No. 1.7e-78;
 Matches 312; Conservative 88; Mismatches 173; Indels 15; Gaps 6;
 16 AVVDCPKYVCVONI;S;E;I;G;T;C;P;S;K;G;L;L;F;V;P;D;I;R;R;V;L;R;L;G;N;F;I;H;I;S;R;O;D;F;A;N;M 75

RESULT 9

; Sequence 100, Application US/10238183
 ; Publication No. US20030073189A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Baker, Kevin
 ; APPLICANT: Eaton, Dan
 ; APPLICANT: Filwaroff, Ellen
 ; APPLICANT: Goddard, Audrey
 ; APPLICANT: Gramaldi, J. Christopher
 ; APPLICANT: Gurney, Austin
 ; APPLICANT: Smith, Victoria
 ; APPLICANT: Stephan, Jean-Phillippe
 ; APPLICANT: Watanabe, Colin
 ; APPLICANT: Wood, William
 ; APPLICANT: Zhang, Zemin
 ; APPLICANT: Fong, Sherman
 ; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
 ; FILE REFERENCE: P3630R1C11
 ; CURRENT APPLICATION NUMBER: US/10/238,183
 ; PRIOR FILING DATE: 2002-09-09
 ; PRIOR APPLICATION NUMBER: 10/197942
 ; PRIOR FILING DATE: 2002-07-18
 ; PRIOR APPLICATION NUMBER: 60/059114
 ; PRIOR FILING DATE: 1997-09-17
 ; PRIOR APPLICATION NUMBER: 60/063046
 ; PRIOR FILING DATE: 1997-10-24
 ; PRIOR APPLICATION NUMBER: 60/065027
 ; PRIOR FILING DATE: 1997-11-10
 ; PRIOR APPLICATION NUMBER: 60/079689
 ; PRIOR FILING DATE: 1998-03-27
 ; PRIOR APPLICATION NUMBER: 60/086478
 ; PRIOR FILING DATE: 1998-05-22

Tue Jun 10 08:57:39 2003

PRIOR APPLICATION NUMBER: 60/087607
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/089801
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/090557
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090689
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/091358
PRIOR FILING DATE: 1998-07-01
PRIOR APPLICATION NUMBER: 60/091978
PRIOR FILING DATE: 1998-07-07
PRIOR APPLICATION NUMBER: 60/099803
PRIOR FILING DATE: 1998-09-10
PRIOR APPLICATION NUMBER: 60/106932
PRIOR FILING DATE: 1998-11-03
PRIOR APPLICATION NUMBER: 60/115554
PRIOR FILING DATE: 1999-01-12
PRIOR APPLICATION NUMBER: 60/119342
PRIOR FILING DATE: 1999-02-09
PRIOR APPLICATION NUMBER: 60/123957
PRIOR FILING DATE: 1999-03-12
PRIOR APPLICATION NUMBER: 60/123972
PRIOR FILING DATE: 1999-03-11
PRIOR APPLICATION NUMBER: 60/127372
PRIOR FILING DATE: 1999-04-01
PRIOR APPLICATION NUMBER: 60/131271
PRIOR FILING DATE: 1999-04-27
PRIOR APPLICATION NUMBER: 60/133459
PRIOR FILING DATE: 1999-05-11
PRIOR APPLICATION NUMBER: 60/135725
PRIOR FILING DATE: 1999-05-25
PRIOR APPLICATION NUMBER: 60/135729
PRIOR FILING DATE: 1999-05-25
PRIOR APPLICATION NUMBER: 60/135750
PRIOR FILING DATE: 1999-05-25
PRIOR APPLICATION NUMBER: 60/138385
PRIOR FILING DATE: 1999-06-09
PRIOR APPLICATION NUMBER: 60/140653
PRIOR FILING DATE: 1999-06-22
PRIOR APPLICATION NUMBER: 60/141037
PRIOR FILING DATE: 1999-06-23
PRIOR APPLICATION NUMBER: 60/144732
PRIOR FILING DATE: 1999-07-20
PRIOR APPLICATION NUMBER: 60/144758
PRIOR FILING DATE: 1999-07-20
PRIOR APPLICATION NUMBER: 60/144790
PRIOR FILING DATE: 1999-07-20
PRIOR APPLICATION NUMBER: 60/145228
PRIOR FILING DATE: 1999-07-20
PRIOR APPLICATION NUMBER: 60/145698
PRIOR FILING DATE: 1999-07-26
PRIOR APPLICATION NUMBER: 60/146222
PRIOR FILING DATE: 1999-07-28
PRIOR APPLICATION NUMBER: 60/146843
PRIOR FILING DATE: 1999-08-03
PRIOR APPLICATION NUMBER: 60/148188
PRIOR FILING DATE: 1999-08-10
PRIOR APPLICATION NUMBER: 60/148513
PRIOR FILING DATE: 1999-08-12
PRIOR APPLICATION NUMBER: 60/149327
PRIOR FILING DATE: 1999-08-17
PRIOR APPLICATION NUMBER: 60/149395
PRIOR FILING DATE: 1999-08-17
PRIOR APPLICATION NUMBER: 60/150114
PRIOR FILING DATE: 1999-08-20
PRIOR APPLICATION NUMBER: 60/151700
PRIOR FILING DATE: 1999-08-31
PRIOR APPLICATION NUMBER: 60/151734
PRIOR FILING DATE: 1999-08-31
PRIOR APPLICATION NUMBER: 60/162506
PRIOR FILING DATE: 1999-10-29
PRIOR APPLICATION NUMBER: 60/170262

PRIOR FILING DATE: 1999-12-09
PRIOR APPLICATION NUMBER: 60/177118
PRIOR FILING DATE: 2000-01-20
PRIOR APPLICATION NUMBER: 60/179851
PRIOR FILING DATE: 2000-02-02
PRIOR APPLICATION NUMBER: 60/180921
PRIOR FILING DATE: 2000-02-08
PRIOR APPLICATION NUMBER: 60/187202
PRIOR FILING DATE: 2000-03-03
PRIOR APPLICATION NUMBER: 60/198587
PRIOR FILING DATE: 2000-04-18
PRIOR APPLICATION NUMBER: 60/199614
PRIOR FILING DATE: 2000-04-25
PRIOR APPLICATION NUMBER: 60/206330
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/206368
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/209832
PRIOR FILING DATE: 2000-06-05
PRIOR APPLICATION NUMBER: 60/218371
PRIOR FILING DATE: 2000-07-13
PRIOR APPLICATION NUMBER: 60/222695
PRIOR FILING DATE: 2000-08-02
PRIOR APPLICATION NUMBER: 60/229896
PRIOR FILING DATE: 2000-09-01
PRIOR APPLICATION NUMBER: 60/230621
PRIOR FILING DATE: 2000-09-05
PRIOR APPLICATION NUMBER: 60/232887
PRIOR FILING DATE: 2000-09-15
PRIOR APPLICATION NUMBER: 60/235147
PRIOR FILING DATE: 2000-09-22
PRIOR APPLICATION NUMBER: 60/261878
PRIOR FILING DATE: 2001-01-12
PRIOR APPLICATION NUMBER: 60/261910
PRIOR FILING DATE: 2001-01-16
PRIOR APPLICATION NUMBER: 60/261939
PRIOR FILING DATE: 2001-01-16
PRIOR APPLICATION NUMBER: 60/262150
PRIOR FILING DATE: 2001-01-16
PRIOR APPLICATION NUMBER: 60/264395
PRIOR FILING DATE: 2001-01-25
PRIOR APPLICATION NUMBER: 60/266421
PRIOR FILING DATE: 2001-02-02
PRIOR APPLICATION NUMBER: 60/267623
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/274399
PRIOR FILING DATE: 2001-03-09
PRIOR APPLICATION NUMBER: 60/280982
PRIOR FILING DATE: 2001-04-03
PRIOR APPLICATION NUMBER: 60/282129
PRIOR FILING DATE: 2001-04-04
PRIOR APPLICATION NUMBER: 60/282199
PRIOR FILING DATE: 2001-04-04
PRIOR APPLICATION NUMBER: 60/290589
PRIOR FILING DATE: 2001-05-09
PRIOR APPLICATION NUMBER: 09/180997
PRIOR FILING DATE: 1998-11-19
PRIOR APPLICATION NUMBER: 09/267213
PRIOR FILING DATE: 1999-03-12
PRIOR APPLICATION NUMBER: 09/380137
PRIOR FILING DATE: 1999-08-25
PRIOR APPLICATION NUMBER: 09/380138
PRIOR FILING DATE: 1999-08-25
PRIOR APPLICATION NUMBER: 09/403297
PRIOR FILING DATE: 1999-10-18
PRIOR APPLICATION NUMBER: 09/423741
PRIOR FILING DATE: 1999-11-10
PRIOR APPLICATION NUMBER: 09/709238
PRIOR FILING DATE: 2000-11-08
PRIOR APPLICATION NUMBER: 09/802706
PRIOR FILING DATE: 2001-03-09
PRIOR APPLICATION NUMBER: 2001-06-01

```

; PRIOR APPLICATION NUMBER: 09/918585
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: 09/924419
; PRIOR FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: 09/927796
; PRIOR FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: 09/929404
; PRIOR FILING DATE: 2001-08-13
; PRIOR APPLICATION NUMBER: 09/931836
; PRIOR FILING DATE: 2001-08-16
; PRIOR APPLICATION NUMBER: 09/941992
; PRIOR FILING DATE: 2001-08-28
; PRIOR APPLICATION NUMBER: 09/946374
; PRIOR FILING DATE: 2001-09-04
; PRIOR APPLICATION NUMBER: 10/001054
; PRIOR FILING DATE: 2001-11-30
; PRIOR APPLICATION NUMBER: 10/052586
; PRIOR FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: 10/081056
; PRIOR FILING DATE: 2002-02-20
; PRIOR APPLICATION NUMBER: 10/119480
; PRIOR FILING DATE: 2002-04-09

```

```

Query Match 38.2%; Score 1563.5; DB 9; Length 627;
Best Local Similarity 53.1%; Pred. No. 1.7e-78; Indels 15; Gaps 6;
Matches 312; Conservative 88; Mismatches 173;

Qy 16 AVVDA CPKYCVQNLSLGLTLCPSKGLLVFPDPIDRRRVLRGKGFIIHISRODFANM 75
Db 23 ATFSPCPRRCRCQTSQSLPLSVLCPCGAGLLFVPPSLDRRAAEURLADNFASFVRRDRLANM 82

Qy 76 TGLVDLTL SRNRTISHIQPFSLDLSRLSLHLDNSRLPSLGBEDTLRGLVNLQHLIVNNQ 135
Db 83 TGLLHLSLRNTRIRVAAGAFADRALRALHLDGNRLTSLGEGQLRGLVNLRLHLSLNQ 142

Qy 136 LGGIADAEDELLTLEDDLSYNLHGLPWSVRMNLHQLSDHNLDDHIAEGTAPD 195
Db 143 LAALAAAGALDCAETLEDLSDSYNNLEQLPWEALGRGNGVNTLGLDHNLLASV-PGAFSR 201

Qy 196 LOKLARLDLTSNRLOKLPDFIFARSOASALATATPPAPLS---FSGGNPLHNCCELLW 252
Db 202 LHKLARLDWTSNRLTIPDPPLFSR---LPLARPRGSPASALVLAFCGNPLHNCCELLW 258

Qy 253 LRRLERDDLETGSGPGLKGRYFVHWREBEFVCEPPLITQHTHKLVLLEGQAATLKCKA 312
Db 259 LRRLAREDDLEACASPPALGGRYFVAWGEFVCEPPVVTTHRSPPPLAVPAGRPAALRCRA 318

Qy 313 IGDPSPLIHWVAPDRDLVGNSSRTAVYDNGTLDIFITTSQDSGAFTCIAANAAGATAMV 372
Db 319 VGDPEPRVRVSPQGRLLGNSRRARAFNGTLLELVTEPDGGIFTFCIAANAAGATAAV 378

Qy 373 EYSI--VQLPPLSNSTRTAPPKSRSLDITGSKTSRGGGGGGGGPPKSPPERAVLVE 430
Db 379 ELTVGPPPPQLANSTSCDPRDGDPAITPASAASAKVADTG---PPTDRGVQVTE 434

Qy 431 VTTTALVKWSKSAVRVMYQOLQVNCSDDEVLIYRMI PASNKAFVNNVNLVSGTYDLC 490
Db 435 HGATAALVQWPOPIRGIMYQIOWNSSADDILVYRMI PAESRSPLTDLASGRTYDLC 494

Qy 491 VLAWDDTATTTATNIVCAOFFKADYPOCOSMHSQILGTMILVIGGIIVATLLVFI 550
Db 495 VLAVYEDSATGTATRPVGCARFSTEPALRPCGAPHPALGTMIALGGMIVASVYLVFI 554

Qy 551 VILMRYKVCNHEAP--SKMAAAVSNVYQNTGAOPPPSPSSAGAPAGAPP 596
Db 555 FVLLMRYKVGQPGKAKIPAVSSVCQNTNGALGFTTTPAPPAP 602

```

```

RESULT 10
US-10-238-283-100
; Sequence 100, Application US/10238283
; Publication No. US20030073190A1
; GENERAL INFORMATION:

```

```

; APPLICANT: Baker, Kevin
; APPLICANT: Eaton, Dan
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Goddard, Audrey
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Phillippe
; APPLICANT: Watanabe, Colin
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; APPLICANT: Fong, Sherman
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3630RIC15
; CURRENT APPLICATION NUMBER: US/10/238,283
; PRIOR FILING DATE: 2002-09-09
; PRIOR APPLICATION NUMBER: 10/197942
; PRIOR FILING DATE: 2002-07-16
; PRIOR APPLICATION NUMBER: 60/059114
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/063046
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/065027
; PRIOR FILING DATE: 1997-11-10
; PRIOR APPLICATION NUMBER: 60/079689
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/086478
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090689
; PRIOR FILING DATE: 1998-06-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 116
; SEQ ID NO 100
; LENGTH: 627
; TYPE: PRT
; ORGANISM: Homo Sapien
; US-10-238-283-100

```

```

Query Match 38.2%; Score 1563.5; DB 9; Length 627;
Best Local Similarity 53.1%; Pred. No. 1.7e-78; Indels 15; Gaps 6;
Matches 312; Conservative 88; Mismatches 173;

Qy 16 AVVDA CPKYCVQNLSLGLTLCPSKGLLVFPDPIDRRRVLRGKGFIIHISRODFANM 75
Db 23 ATFSPCPRRCRCQTSQSLPLSVLCPCGAGLLFVPPSLDRRAAEURLADNFASFVRRDRLANM 82

Qy 76 TGLVDLTL SRNRTISHIQPFSLDLSRLSLHLDNSRLPSLGBEDTLRGLVNLQHLIVNNQ 135
Db 83 TGLLHLSLRNTRIRVAAGAFADRALRALHLDGNRLTSLGEGQLRGLVNLRLHLSLNQ 142

Qy 136 LGGIADAEDELLTLEDDLSYNLHGLPWSVRMNLHQLSDHNLDDHIAEGTAPD 195
Db 143 LAALAAAGALDCAETLEDLSDSYNNLEQLPWEALGRGNGVNTLGLDHNLLASV-PGAFSR 201

Qy 196 LOKLARLDLTSNRLOKLPDFIFARSOASALATATPPAPLS---FSGGNPLHNCCELLW 252
Db 202 LHKLARLDWTSNRLTIPDPPLFSR---LPLARPRGSPASALVLAFCGNPLHNCCELLW 258

Qy 253 LRRLERDDLETGSGPGLKGRYFVHWREBEFVCEPPLITQHTHKLVLLEGQAATLKCKA 312
Db 259 LRRLAREDDLEACASPPALGGRYFVAWGEFVCEPPVVTTHRSPPPLAVPAGRPAALRCRA 318

Qy 313 IGDPSPLIHWVAPDRDLVGNSSRTAVYDNGTLDIFITTSQDSGAFTCIAANAAGATAMV 372
Db 319 VGDPEPRVRVSPQGRLLGNSRRARAFNGTLLELVTEPDGGIFTFCIAANAAGATAAV 378

```

QY 373 EVSI--VOLPHLSNSTRTAPPKSRSLDITGSSKTSRGGGSGGCGEPPKSPPERAVLYSE 430
 Db 379 ELTVGPPPPQLANSTSCDPPRDGDDALTPPSAASASAKVADTG---PPTDRGVQVTE 434
 QY 431 VTTTSALVKWSKSAAPRVKMYQLOVNCSDDEVLIVYRMIIPASNKAFVNNLVSGTGYDLC 490
 Db 435 HGATAALVQWPDQRPPIGIRMYQIQYNSSADDILVYRMIIPAESRSFLLTDLASGRTYDLC 494
 QY 491 VLAMWDDTATLTATNIVGCAOFFTKADYPOCOSMHSOILGGTMIIVGGIIVATLLVFI 550
 Db 495 VLAVYEDSATGLTATRPVGCARFSTEPALRPCGAPHAPFLGTTMIIALGGVIVASVLVFI 554
 QY 551 VILMRYKYCNHEAP--SKMAAAVSNVYSGTNGAOPPPSSAPAGAPP 596
 Db 555 FVLLMRYKHGGQPPGKAKIPAPVSSVCSQTNALGPTTTPAPPAP 602

RESULT 11
 US-10-238-370-100
 ; Sequence 100, Application US/10238370
 ; Publication No. US20030073191A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Baker, Kevin
 ; APPLICANT: Eaton, Dan
 ; APPLICANT: Filvaroff, Ellen
 ; APPLICANT: Goddard, Audrey
 ; APPLICANT: Grimaldi, J. Christopher
 ; APPLICANT: Gurney, Austin
 ; APPLICANT: Smith, Victoria
 ; APPLICANT: Stephan, Jean-Phillippe
 ; APPLICANT: Watambe, Colin
 ; APPLICANT: Wood, William
 ; APPLICANT: Zhang, Zemin
 ; APPLICANT: Fong, Sherman
 ; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
 ; FILE REFERENCE: P3630R1C10
 ; CURRENT APPLICATION NUMBER: US/10/238,370
 ; CURRENT FILING DATE: 2002-09-09
 ; PRIOR APPLICATION NUMBER: 10/197942
 ; PRIOR FILING DATE: 2002-07-18
 ; PRIOR APPLICATION NUMBER: 60/059114
 ; PRIOR FILING DATE: 1997-09-17
 ; PRIOR APPLICATION NUMBER: 60/063046
 ; PRIOR FILING DATE: 1997-10-24
 ; PRIOR APPLICATION NUMBER: 60/065027
 ; PRIOR FILING DATE: 1997-11-10
 ; PRIOR APPLICATION NUMBER: 60/079689
 ; PRIOR FILING DATE: 1998-03-27
 ; PRIOR APPLICATION NUMBER: 60/086478
 ; PRIOR FILING DATE: 1998-05-22
 ; PRIOR APPLICATION NUMBER: 60/087607
 ; PRIOR FILING DATE: 1998-06-02
 ; PRIOR APPLICATION NUMBER: 60/089801
 ; PRIOR FILING DATE: 1998-06-18
 ; PRIOR APPLICATION NUMBER: 60/090557
 ; PRIOR FILING DATE: 1998-06-24
 ; PRIOR APPLICATION NUMBER: 60/090689
 ; PRIOR FILING DATE: 1998-06-25
 ; Remaining Prior Application data removed - See File Wrapper or PALM.
 ; NUMBER OF SEQ ID NOS: 116
 ; SEQ ID NO 100
 ; LENGTH: 627
 ; TYPE: PRY
 ; ORGANISM: Homo Sapien
 US-10-238-370-100

Db 23 ATPSPCRRRCQTSQLEPLSVLPCFAGLLFVPPSLDRRAAEELRLADNFIAVRRRDLANM 82
 QY 76 TGLVDLTLRNTI SHIQPFPLDLESRLHLDNSNRLPSLGEDTLRGLVNLQHLIVNNQ 135
 Db 83 TGLUHLSLSRNTIRHVAGAFADLRALRALHLDGNRLTSLGEGQLRGLVNLRHLLSNQ 142
 QY 136 LGGIADAEFEDLTLLEDDLSYNNLHGLPMDSVRRMVNLHQLSLDHNHLLDHIAGEFTAD 195
 Db 143 LAALAGALDDCAETLEDLSDSYNNLQPLPWEALGRGLNVTNLGLDHNHLLASV-PGAFSR 201
 QY 196 LOKLARLDLTSNRLOKLPDPPIPARSOASALTATFPAPPLS---FSGGNPLHNCCELLW 252
 Db 202 LHKLARLDMTSNRLTTPDPLFSR---LPLLARPRGSPASALVLAFCGNPLHNCCELLW 258
 QY 253 LRRLEDDLETCSGGKGRYFVHWVREBEFVCEPPLITQHTHKLVLLEGOAATLKCKA 312
 Db 259 LRRLAREDDLEACASPPALGGRYFWAYGEEFVCEPVPVWTHRSPPPLAVPAGRPAALRCEA 318
 QY 313 IGDPSLIHWVAPDDRLLVGNSSRTAVYDNGTLDIFITTSODSGAFTCIAANAAGEATAMV 372
 Db 319 VGDPEPRVWVSPQGRLLGNSSRARAFPNGTLELLVTEPPGDDGIFTCIAANAAGEATAAV 378
 QY 373 EVSI--VOLPHLSNSTRTAPPKSRSLDITGSSKTSRGGGSGGCGEPPKSPPERAVLYSE 430
 Db 379 ELTVGPPPPQLANSTSCDPPRDGDDALTPPSAASASAKVADTG---PPTDRGVQVTE 434
 QY 431 VTTTSALVKWSKSAAPRVKMYQLOVNCSDDEVLIVYRMIIPASNKAFVNNLVSGTGYDLC 490
 Db 435 HGATAALVQWPDQRPPIGIRMYQIQYNSSADDILVYRMIIPAESRSFLLTDLASGRTYDLC 494
 QY 491 VLAMWDDTATLTATNIVGCAOFFTKADYPOCOSMHSOILGGTMIIVGGIIVATLLVFI 550
 Db 495 VLAVYEDSATGLTATRPVGCARFSTEPALRPCGAPHAPFLGTTMIIALGGVIVASVLVFI 554
 QY 551 VILMRYKYCNHEAP--SKMAAAVSNVYSGTNGAOPPPSSAPAGAPP 596
 Db 555 FVLLMRYKHGGQPPGKAKIPAPVSSVCSQTNALGPTTTPAPPAP 602

RESULT 12
 US-10-245-055-100
 ; Sequence 100, Application US/10245055
 ; Publication No. US20030073192A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Baker, Kevin
 ; APPLICANT: Eaton, Dan
 ; APPLICANT: Filvaroff, Ellen
 ; APPLICANT: Goddard, Audrey
 ; APPLICANT: Grimaldi, J. Christopher
 ; APPLICANT: Gurney, Austin
 ; APPLICANT: Smith, Victoria
 ; APPLICANT: Stephan, Jean-Phillippe
 ; APPLICANT: Watambe, Colin
 ; APPLICANT: Wood, William
 ; APPLICANT: Zhang, Zemin
 ; APPLICANT: Fong, Sherman
 ; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
 ; FILE REFERENCE: P3630R1C88
 ; CURRENT APPLICATION NUMBER: US/10/245,055
 ; CURRENT FILING DATE: 2002-09-16
 ; PRIOR APPLICATION NUMBER: 10/197942
 ; PRIOR FILING DATE: 2002-07-18
 ; PRIOR APPLICATION NUMBER: 60/059114
 ; PRIOR FILING DATE: 1997-09-17
 ; PRIOR APPLICATION NUMBER: 60/063046
 ; PRIOR FILING DATE: 1997-10-24
 ; PRIOR APPLICATION NUMBER: 60/065027
 ; PRIOR FILING DATE: 1997-11-10
 ; PRIOR APPLICATION NUMBER: 60/079689
 ; PRIOR FILING DATE: 1998-03-27
 ; PRIOR APPLICATION NUMBER: 60/086478
 ; PRIOR FILING DATE: 1998-05-22

Query Match 38.2%; Score 1563.5; DB 9; Length 627;
 Best Local Similarity 53.1%; Pred. No. 1.7e-78;
 Matches 312; Conservative 173; Indels 15; Gaps 6;
 QY 16 AVWDAPKYCVQNLSSIGTLCPKGLLFFVPPDDIRRTVELLIGGNFIHISRODFANM 75


```

; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090689
; PRIOR FILING DATE: 1998-06-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 116
; SEQ ID NO 100
; LENGTH: 627
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-245-055-100

```

```

Query Match 38.2%; Score 1563.5; DB 9; Length 627;
Best Local Similarity 53.1%; Pred. No. 1.7e-78;
Matches 312; Conservative 88; Mismatches 173; Indels 15; Gaps 6;

```

```

QY 16 AVVDAAPKYVCQNLSESLGTLCPKSGKLLFVPPDDIRRTVELRGLGNFIHISRQDFAMN 75
  |||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||
Db 23 ATPSPRRRCRCQOTQSLPLVPCGAGLLFVPPSLDRRAAEELRDLADNFIASVRRDRLANM 82
  |||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||
QY 76 TGLVDLTLRSNTISHIQPPFDLLESRLSDNSRLSLGEDTLRGLVNLQHLVNNQ 135
  |||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||
Db 83 TGLLHLSLRNTRIRHVAAGAFADRALRALHLDGRLNRLTSLEGGQLRGLVNLHRLHLSNNQ 142
  |||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||
QY 136 LGGIADAEFEDLLTLEDLSDSYNNLHGLPWDSVRRMVNLHOLSLDHLLDDHIAEGETAD 195
  |||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||
Db 143 LAALAGALDDCAETLEDDLSDSYNNLHGLPWDSVRRMVNLHOLSLDHLLDDHIAEGETAD 201
  |||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||
QY 196 LOKLARLDLTLNRLKLPDPPIFARSQASALATATPPAPPLS---FSFGNPLHNCCELLW 252
  |||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||
Db 202 LHKLARLDWTSNRLTTPDPFLFSR---LPLLARPGSPASALVLAFCGNPLHNCCELLW 258
  |||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||
QY 253 LRRLERDDELTCGSPGLKGRYFHWVREBEEFVCEPPLITQHTKLLVLEGGQAATLKCKA 312
  |||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||
Db 259 LRRLAREDDLEACASPPALGGRYFWAVGEEFVCEPPVVTHRSPPPLAVPAGRPAALRCRA 318
  |||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||
QY 313 IGDRPSPLHVAAPDRDRLVGNSSRTAVYDNGTLDIFITTSQDSGAFCTIAANAAGEATAMV 372
  |||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||
Db 319 VGDPEPRVWVSPQGRLLGNSRRARAFNGTLELIVTEPDGGGIFTCIAANAAGEATRAV 378
  |||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||
QY 373 EIVI--VOLPHLSNSTRTAPPKSRSLSDITGSSKTSRGGGSGGEPKSPPERAVLYSE 430
  |||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||
Db 379 ELTVGPPPPPQLANSTSCDPPRDGPDALTTPSAASASAKVADTG---PPTDRGVQVTE 434
  |||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||
QY 431 VTTSALVWKSVSKAPRVKMYQOYNCSDDDEVLYRMIIPASNKAFFVNNLVSGTGFDLC 490
  |||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||
Db 435 HGATAALVQWPDPPIPIRMYQIYNSASADLIVRMIPAESRSFLLTDLASGRTYDLC 494
  |||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||
QY 491 VLAMWDDTTATLTATNIVCAOFFKADYPCQSMHSQLCGTMILVIGGIIVATLLVPI 550
  |||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||
Db 495 VLAVYEDSATSATATRVGCARFTEPALRPCGAPHAPLGGTMIALGGVIVASVLVPI 554
  |||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||
QY 551 VLLMRYKYVGNHEAP--SKMAAANSVYVOTNGAQQPPPPSSAPACAPP 596
  |||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||
Db 555 FVLLMRYKYVHGQPCKAKIAPVSVCSQTMALGPTPPAPPEP 602
  |||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||

```

```

RESULT 13
US-10-245-147-100
; Sequence 100, Application US/10245147
; Publication No. US20030073193A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin
; APPLICANT: Eaton, Dan
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Goddard, Audrey
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin
; APPLICANT: Smith, Victoria

```

```

; APPLICANT: Stephan, Jean-Phillippe
; APPLICANT: Watanabe, Colin
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; APPLICANT: Fong, Sherman
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3630RIC72
; CURRENT APPLICATION NUMBER: US/10/245.147
; CURRENT FILING DATE: 2002-09-16
; PRIOR APPLICATION NUMBER: 10/197942
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/059114
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/063046
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/065027
; PRIOR FILING DATE: 1997-11-10
; PRIOR APPLICATION NUMBER: 60/079689
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/086478
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090689
; PRIOR FILING DATE: 1998-06-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 116
; SEQ ID NO 100
; LENGTH: 627
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-245-147-100

```

```

Query Match 38.2%; Score 1563.5; DB 9; Length 627;
Best Local Similarity 53.1%; Pred. No. 1.7e-78;
Matches 312; Conservative 88; Mismatches 173; Indels 15; Gaps 6;

```

```

QY 16 AVVDAAPKYVCQNLSESLGTLCPKSGKLLFVPPDDIRRTVELRGLGNFIHISRQDFAMN 75
  |||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||
Db 23 ATPSPRRRCRCQOTQSLPLVPCGAGLLFVPPSLDRRAAEELRDLADNFIASVRRDRLANM 82
  |||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||
QY 76 TGLVDLTLRSNTISHIQPPFDLLESRLSDNSRLSLGEDTLRGLVNLQHLVNNQ 135
  |||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||
Db 83 TGLLHLSLRNTRIRHVAAGAFADRALRALHLDGRLNRLTSLEGGQLRGLVNLHRLHLSNNQ 142
  |||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||
QY 136 LGGIADAEFEDLLTLEDLSDSYNNLHGLPWDSVRRMVNLHOLSLDHLLDDHIAEGETAD 195
  |||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||
Db 143 LAALAGALDDCAETLEDDLSDSYNNLHGLPWDSVRRMVNLHOLSLDHLLDDHIAEGETAD 201
  |||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||
QY 196 LOKLARLDLTLNRLKLPDPPIFARSQASALATATPPAPPLS---FSFGNPLHNCCELLW 252
  |||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||
Db 202 LHKLARLDWTSNRLTTPDPFLFSR---LPLLARPGSPASALVLAFCGNPLHNCCELLW 258
  |||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||
QY 253 LRRLERDDELTCGSPGLKGRYFHWVREBEEFVCEPPLITQHTKLLVLEGGQAATLKCKA 312
  |||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||
Db 259 LRRLAREDDLEACASPPALGGRYFWAVGEEFVCEPPVVTHRSPPPLAVPAGRPAALRCRA 318
  |||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||
QY 313 IGDRPSPLHVAAPDRDRLVGNSSRTAVYDNGTLDIFITTSQDSGAFCTIAANAAGEATAMV 372
  |||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||
Db 319 VGDPEPRVWVSPQGRLLGNSRRARAFNGTLELIVTEPDGGGIFTCIAANAAGEATRAV 378
  |||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||
QY 373 EIVI--VOLPHLSNSTRTAPPKSRSLSDITGSSKTSRGGGSGGEPKSPPERAVLYSE 430
  |||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||
Db 379 ELTVGPPPPPQLANSTSCDPPRDGPDALTTPSAASASAKVADTG---PPTDRGVQVTE 434
  |||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||
QY 431 VTTSALVWKSVSKAPRVKMYQOYNCSDDDEVLYRMIIPASNKAFFVNNLVSGTGFDLC 490
  |||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||
Db 435 HGATAALVQWPDPPIPIRMYQIYNSASADLIVRMIPAESRSFLLTDLASGRTYDLC 494
  |||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||
QY 491 VLAMWDDTTATLTATNIVCAOFFKADYPCQSMHSQLCGTMILVIGGIIVATLLVPI 550
  |||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||
Db 495 VLAVYEDSATSATATRVGCARFTEPALRPCGAPHAPLGGTMIALGGVIVASVLVPI 554
  |||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||
QY 551 VLLMRYKYVGNHEAP--SKMAAANSVYVOTNGAQQPPPPSSAPACAPP 596
  |||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||
Db 555 FVLLMRYKYVHGQPCKAKIAPVSVCSQTMALGPTPPAPPEP 602
  |||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||

```

```

RESULT 13
US-10-245-147-100
; Sequence 100, Application US/10245147
; Publication No. US20030073193A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin
; APPLICANT: Eaton, Dan
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Goddard, Audrey
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin
; APPLICANT: Smith, Victoria

```

QY 491 VLAMDDTATLTATNIVGCAOFFTKADYPOCOSMHSQILGGTMIIVIGIIVATLLVFI 550
 Db 495 VLVAVDSATGUTATRPVGCARFSTPALRPCGAPHPFLGGTMIILGGLVIVASVLVFI 554
 QY 551 VILVMRYKVCNHEAP--SKMAAAVSNVYQTNGAQPPSSAPAGAPP 596
 Db 555 FVLLMRYKHGGQPPGKAKIPAPVSSVCSQTNALGPTTPAPPAPPEP 602

RESULT 14
 US-10-245-730-100
 ; Sequence 100, Application US/10245730
 ; Publication No. US20030073194A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Baker, Kevin
 ; APPLICANT: Eaton, Dan
 ; APPLICANT: Filvaroff, Ellen
 ; APPLICANT: Goddard, Audrey
 ; APPLICANT: Grimaldi, J. Christopher
 ; APPLICANT: Gurney, Austin
 ; APPLICANT: Smith, Victoria
 ; APPLICANT: Stephan, Jean-Phillippe
 ; APPLICANT: Watanabe, Colin
 ; APPLICANT: Wood, William
 ; APPLICANT: Zhang, Zemin
 ; APPLICANT: Fong, Sherman
 ; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
 ; TITLE OF INVENTION: ACIDS ENCODING THE SAME
 ; FILE REFERENCE: P3630R1C85
 ; CURRENT APPLICATION NUMBER: US/10/245,730
 ; CURRENT FILING DATE: 2002-09-16
 ; PRIOR APPLICATION NUMBER: 10/197942
 ; PRIOR FILING DATE: 2002-07-18
 ; PRIOR APPLICATION NUMBER: 60/059114
 ; PRIOR FILING DATE: 1997-09-17
 ; PRIOR APPLICATION NUMBER: 60/063046
 ; PRIOR FILING DATE: 1997-10-24
 ; PRIOR APPLICATION NUMBER: 60/065027
 ; PRIOR FILING DATE: 1997-11-10
 ; PRIOR APPLICATION NUMBER: 60/079689
 ; PRIOR FILING DATE: 1998-03-27
 ; PRIOR APPLICATION NUMBER: 60/086478
 ; PRIOR FILING DATE: 1998-05-22
 ; PRIOR APPLICATION NUMBER: 60/087607
 ; PRIOR FILING DATE: 1998-06-02
 ; PRIOR APPLICATION NUMBER: 60/089801
 ; PRIOR FILING DATE: 1998-06-18
 ; PRIOR APPLICATION NUMBER: 60/090557
 ; PRIOR FILING DATE: 1998-06-24
 ; PRIOR APPLICATION NUMBER: 60/090689
 ; PRIOR FILING DATE: 1998-06-25
 ; Remaining prior Application data removed - See File Wrapper or PALM.
 ; NUMBER OF SEQ ID NOS: 116
 ; SEQ ID NO 100
 ; LENGTH: 627
 ; TYPE: PRT
 ; ORGANISM: Homo Sapien
 ; US-10-245-730-100

Query Match 38.2%; Score 1563.5; DB 9; Length 627;
 Best Local Similarity 53.1%; Pred. No. 1.7e-78;
 Matches 312; Conservative 89; Mismatches 173; Indels 15; Gaps 6;
 QY 16 AVVDAPKVCVQNLSESLGTLCPKGLLFPVPPDIRTVLELKGNFIIHISRQDFANM 75
 Db 23 ATPSPCPRCKCQTSQSLPLSVLFCGAGLLFPVPSLDRRAELADNFIAFVRRRLANM 82
 QY 76 TGLVDTLTSRNTLSHQPFQFLESLRSJLHDSNRLPSJGEDTLRGLVNLQHLIWNNO 135
 Db 83 TGLHLSLSRNTLRHVNAFADRALRHLHDSNRLTSJGEGQLRGLVNLRLHLSNQ 142
 QY 136 LGGIADAEFDFLLTLEDLDSYNNLHGLPWDSVRRVNLHQLSLDHNLLDHTAEGTFAD 195

Db 143 LAALAAAGALDDCAETLEDLDSYNNLEQLPWEALGRGNVNTLGDHNLASV-PGAFSR 201
 QY 196 LOKLARLDLTSNRLOKLPDPDFARSOASALTATPFAPPLS---FSGNPLHCNCELLW 252
 Db 202 LHKLARLDMTSNRLTTPPPDFLFSR---LPLARPRGSPASALVAFGNGNPLHCNCELLW 258
 QY 253 LRLERDDDLTCCSPGGLKGRYFHWVREBEFVCEPPLIQTHTKLLVLEGOAATLKCKA 312
 Db 259 LRLAREDDLEACASPPALGGRYFWAYGEEFVCEPVPVTHRSPPPLAVPAGRALRCRA 318
 QY 313 IGDEPPLIHWVAPDDRLLVGNSSRTAVYDNGTLDIFITTSQDSGAFPTCIAANAAGBATMV 372
 Db 319 VGDPEPRVRVWVSPQCRLLGNSSRARAFPNGLTELLVTEPPGGGIFTCIAANAAGATAAV 378
 QY 373 EVSI--VQLPHLSNSTRTAPPKSRLSDI TCGSSKXTRSGGGSGGEGPEPKSPERAVLYSE 430
 Db 379 ELTGVPPPPQLANSTSCDPPRDPDALTTPSRAASAKVADTG---PPTDRGVQVTE 434
 QY 431 VTTTSALVKWSVKSAPRVKMYQLQYNGSDDEVLIRYMI PASNKAFVYVNNLVSGTGYDLC 490
 Db 435 HGATAALVQWPDQRPICGIRMYQIQYNSADDILVYRMI PAESRSFLLTDLASGRTYDLC 494
 QY 491 VLAMDDTATLTATNIVGCAOFFTKADYPOCOSMHSQILGGTMIIVIGIIVATLLVFI 550
 Db 495 VLVAVDSATGUTATRPVGCARFSTPALRPCGAPHPFLGGTMIILGGLVIVASVLVFI 554
 QY 551 VILVMRYKVCNHEAP--SKMAAAVSNVYQTNGAQPPSSAPAGAPP 596
 Db 555 FVLLMRYKHGGQPPGKAKIPAPVSSVCSQTNALGPTTPAPPAPPEP 602

RESULT 15
 US-10-245-739-100
 ; Sequence 100, Application US/10245739
 ; Publication No. US20030073195A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Baker, Kevin
 ; APPLICANT: Eaton, Dan
 ; APPLICANT: Filvaroff, Ellen
 ; APPLICANT: Goddard, Audrey
 ; APPLICANT: Grimaldi, J. Christopher
 ; APPLICANT: Gurney, Austin
 ; APPLICANT: Smith, Victoria
 ; APPLICANT: Stephan, Jean-Phillippe
 ; APPLICANT: Watanabe, Colin
 ; APPLICANT: Wood, William
 ; APPLICANT: Zhang, Zemin
 ; APPLICANT: Fong, Sherman
 ; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
 ; TITLE OF INVENTION: ACIDS ENCODING THE SAME
 ; FILE REFERENCE: P3630R1C86
 ; CURRENT APPLICATION NUMBER: US/10/245,739
 ; CURRENT FILING DATE: 2002-09-16
 ; PRIOR APPLICATION NUMBER: 10/197942
 ; PRIOR FILING DATE: 2002-07-18
 ; PRIOR APPLICATION NUMBER: 60/059114
 ; PRIOR FILING DATE: 1997-09-17
 ; PRIOR APPLICATION NUMBER: 60/063046
 ; PRIOR FILING DATE: 1997-10-24
 ; PRIOR APPLICATION NUMBER: 60/065027
 ; PRIOR FILING DATE: 1997-11-10
 ; PRIOR APPLICATION NUMBER: 60/079689
 ; PRIOR FILING DATE: 1998-03-27
 ; PRIOR APPLICATION NUMBER: 60/086478
 ; PRIOR FILING DATE: 1998-05-22
 ; PRIOR APPLICATION NUMBER: 60/087607
 ; PRIOR FILING DATE: 1998-06-02
 ; PRIOR APPLICATION NUMBER: 60/089801
 ; PRIOR FILING DATE: 1998-06-18
 ; PRIOR APPLICATION NUMBER: 60/090557
 ; PRIOR FILING DATE: 1998-06-24
 ; PRIOR APPLICATION NUMBER: 60/090689

; PRIOR FILING DATE: 1998-06-25
 ; Remaining Prior Application data removed - See File Wrapper or PALM.
 ; NUMBER OF SEQ ID NOS: 116
 ; SEQ ID NO 100
 ; LENGTH: 627
 ; TYPE: PR1
 ; ORGANISM: Homo Sapien
 ; US-10-245-739-100

Query Match 38.2%; Score 1563.5; DB 9; Length 627;
 Best Local Similarity 53.1%; Pred. No. 1.7e-78;
 Matches 312; Conservative 88; Mismatches 173; Indels 15; Gaps 6;

```

QY 16 AVVDACPKYVCVQNLSLSEGLTICPSKGLLVFPDPIDRRIVVELRLGNGFIIHISRQDFANN 75
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 23 ATPSPCPRCRCQOTQSLPFLVCPGAGLLFVPELDRRAAEELRDLADNFASFVRRRDLANN 82
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 76 TGLVDLTLNRNTHIOPFSLDLESRLSHLSDNRLPSLGEDTLRGLVNLQHLIYVNNQ 135
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 83 TGLLHLSLRNTHIRHVAAAGAFADRALRALHLDGNRLTSLGEGQLRGLVNLRLHLLISNNQ 142
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 136 LGGIADPEAFEDLITLEDLSDLYNNLHGLPWSVRVMNVLHQLSLDHNLLDHIAGCTFAD 195
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 143 LAALAAGALDDCAETLEDLSDLYNNLEQLPWEALGRUGNVNTLGLDHNLLASV-PCAFSR 201
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 196 LOKLARLDLTSNRLOKLPDPPIPARSOASALTATPPAPPLS---FSFGNPLHCNCELLW 252
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 202 LHKLARLDMTSNRLTTPPPDLFSR--LPLLARPRGSPASALVLAFCGNGPLHCNCELLW 258
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 253 LARLEDDDLLETGSPGGLKGRYFWHVRSEEFYCEPPLITQHTHKLVLLEGQATLKCKA 312
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 259 LRRLAREDDLEACASPPALGGRYFWAVGBEEFVCEPPVTVHRSPPLAVPAGRPAALRCRA 318
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 313 IGDPPSLIHWAPDDDLVGNSSRTAVYDNGTLDIFITTSODSGAFTCIAANAAGEATAMV 372
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 319 VGDPEPRVAVVSPQGRLLGNSSRARAFPNGTLELLVTEFGDGGIFTCIAANAAGEATAV 378
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 373 EVSI--VQLPHLSNSTRTAPPKRSLSIDITGSSKTSRGGGGGGGPPKSPPERAVLVSE 430
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 379 ELTVGPPPPPOLANSTSCDPPRDGDPDALTPTPSAASAKVADTG---PPTDRGVQVTE 434
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 431 VTTTSALVWMSVSKSAPRVKMYQVNCSDDEVLIIYRMI PASNKAFVWVNNLYSGTGYDLG 490
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 435 HGATAALVQMPDQRPPIGRIMYQIQYNSADDILVYRMI PAESRSELLTDLASGRTYDLG 494
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 491 VLAMMDTATTTLATNIWGCAOFFTKADYPOCOSHSQILGGTMIIVGIIIVATLLVFI 550
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 495 VLAIVEDSATGLTATRPVGCARSTEPALPCGAPHAPFLGGTMIIALGGVIVASVIVFI 554
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 551 VILMRYKYCNHEAP--SKMAAAVSNVYQTNGAQPPPPSSAPAGAPP 596
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 555 FVLLMRYKVHGQPQPKAKIPAPVSSVCSQTNALGPTPTPAPPAPPEP 602
  
```

Search completed: June 8, 2003, 21:38:55
 Job time : 59 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 8, 2003, 21:20:53 ; Search time 44 Seconds
(without alignments)
1723.865 Million cell updates/sec

Title: US-09-831-846-2
Perfect score: 4094
Sequence: 1 METLLGGLLAFGMFAVVDA.....DLVGARFTFGSSEWVMSTV 789
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues 283224
Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 73.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Table with 5 columns: Result No., Score, Query Match, Length, DB ID, Description. Contains 29 entries of search results.

30 197 4.8 560 2 A60164 platelet membrane
31 196 4.8 821 1 A39667 brain-derived neur
32 195.5 4.8 458 2 T19941 hypothetical prote
33 195.5 4.8 1344 2 T14316 rig-1 protein - mo
34 195 4.8 360 2 S06280 decorin precursor
35 195 4.8 1097 2 A29943 Toll protein precu
36 193 4.7 562 2 T34319 hypothetical prote
37 193 4.7 653 2 T25194 hypothetical prote
38 191 4.7 702 2 T21148 hypothetical prote
39 190.5 4.7 961 2 T23395 hypothetical prote
40 189.5 4.6 476 1 A35104 brain-derived neur
41 188.5 4.6 1257 2 A88536 protein B0523.5 [i
42 186.5 4.6 375 2 S05390 fibromodulin precu
43 186.5 4.6 1134 1 A29944 chaptin precursor
44 185.5 4.5 354 2 A54454 decorin precursor
45 185.5 4.5 369 2 S20811 proteoglycan I - m

ALIGNMENTS

RESULT 1
JC7763
neuronal leucine-rich repeat protein-3 - rat
C:Species: Rattus norvegicus (Norway rat)
C>Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 01-Feb-2002
C:Accession: JC7763
R:Fukamachi, K.; Matsuo, Y.; Kitano, C.; Kuchino, Y.; Tsuda, H.
Biochem. Biophys. Res. Commun. 287, 257-263, 2001
A>Title: Rat neuronal leucine-rich repeat protein-3: Cloning and regulation of the gene
A:Reference number: JC7763; PMID:11549284
A:Contents: Fibrosarcoma cells
A:Accession: JC7763
A:Molecule type: mRNA
A:Residues: 1-707 <FUK>
A:Cross-references: GB:A291437
C:Comment: This protein, a new member of the neuronal leucine-rich repeat protein family
in protein-protein interaction and functions as a cell adhesion molecule or soluble liga
C:Genetics:
A:Gene: nlr-3
C:Keywords: cell adhesion

Query Match 9.3%; Score 380.5; DB 2; Length 707;
Best Local Similarity 22.4%; Pred. No. 3.5e-17;
Matches 154; Conservative 97; Mismatches 233; Indels 203; Gaps 27;
QY 12 GMAF-AVYDA-----CPKYVCQ-----NLSESLGTL-CPKGLLVFPDPDRRTV 55
Db 13 GLAITALVQAGDKKVDPCPOLCTCEIRPWFTEPRSYMEASTVDCNDLGLLNFARLPADTQ 72
QY 56 ELRLGGNFIIHSRQ-DF-ANMTGLVDLTLRSNTISHIQPFSLDLESLSRSLHLSNRLP 113
Db 73 ILLLTQNNIARIEHSTDFPVNLG---LDLSQNNLSSVNTINVKMSQLLSVLEENKLT 129
QY 114 SIGEDTLRGLVNLQHLIYNNNLGGIADFAE-----EDFLTLLEDLD-LSY--NNLH 162
Db 130 ELPERCLYGLSLOELYVHNLNLLSAISPGAFVGLHNLRLHLSNRLQMINSKWEALPN 189
QY 145 -----EDFLTLLEDLD-LSY--NNLH 162
Db 190 LEILMLGNDPIRLIKMDNFQPLLKURSLVIAGINLTFVDDALVGLNLESISFYDNRUN 249
QY 163 GLPMDSVRRMNVNLHSLDHLNLLDHIAGTFA-----DIQ 197
Db 250 KYPVALQKAVNLKFLDLNKNPINIRGDFSNMLHLKELGNNMPELVISLAVDNL 309
QY 198 KLARLDLTSN-RLQKLPDPPIPARSOASALTATPPA-----PPL-SFSGGNP 243
Db 310 DLRKIEATNPRSLYIHPNAFRLPKLESMLNSNALGALYHGTIESLPNLKEISHSNP 369
QY 244 LHCNCELLWLR-----RLERDDDLLETCSFGGLKGRYPWHVREEF-----VCPPLIT 292
Db 370 IRCDVIRNNKNTNIRFMPEPDSL-FCVDPPEFOGQ----NVRQVHFRDMMEICLPLIAP 425

QY 293 QHTHLLVLEGOA-ATLKCKAIGDPSPLHWPAD-DRLVGNSRRTA--VYDNGTLDLFI 348
 Db 426 ESFPSILDVEADYSVSLHCRATAEQPELYWITPSSGKRLLEPNTLREKFKVHSEGLDARG 485
 QY 349 TTSQDSGAFTCIAANAAGEATAMVEVSIQVLPKLSNSTRTAPPKSRSLSDITGSSKTSRG 408
 Db 486 ITPKEGGLYTCIATNLVGDALASIMIKV----- 513
 QY 409 GGSVGGGPPKPPERAVLVSEVTTTSALVKSIV-SKSAPRVKMLQVNCSDDEVLYR 467
 Db 514 ----GGFVPODNGNSLNIKIRDIRANSVLVSKANSKILKSSVKWTFVFKTSDSQAQSA 569
 QY 468 MIPASNKAFVNNLVSGTYDVCVLAWDDTATLTATNIVCAQFFTKA---DYPOCOS 524
 Db 570 RIPSDDVKNVNLTHLKPSTBYKICI-----DIPTIYQSRKOCVNVVTKSLSEHDKENGK 623
 QY 525 MHSQILGGTMMILVIGI--IVATLVF 549
 Db 624 SHT-----VFVACVGGLLIGVMCLF 645
 RESULT 2
 A:Species: Mus musculus (house mouse)
 C:Date: 11-Apr-1997 #sequence_revision 11-Apr-1997 #text_change 05-Nov-1999
 C:Accession: A58532
 R:Suzuki, Y.; Sato, N.; Tohyama, M.; Wanaka, A.; Takagi, T.
 J. Biol. Chem. 271, 22522-22527, 1996
 A:Title: cDNA cloning of a novel membrane glycoprotein that is expressed specifically in
 A:Reference number: A58532, MUID:96394313, PMID:8798419
 A:Accession: A58532
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-1091 <SUZ>
 A:Cross-references: GB:D78572; NID:g1545806; PID:BA11416.1; PID:g1545807
 C:Superfamily: leucine-rich alpha-2-glycoprotein repeat homology; proteoglycan amino-ter
 F:36-61/Domain: proteoglycan amino-terminal homology <PAH>
 F:71-94/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR1>
 F:95-117/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR2>
 F:118-141/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR3>
 F:142-165/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR4>
 F:166-189/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR5>
 F:191-213/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR6>
 F:214-237/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR7>
 F:238-261/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR8>
 F:262-285/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR9>
 F:286-309/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR10>
 F:310-333/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR11>
 F:334-357/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR12>
 F:358-381/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR13>
 F:385-408/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR14>
 F:409-432/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR15>
 F:440-485/Domain: proteoglycan carboxyl-terminal homology <PCH>
 Query Match 9.0%; Score 368; DB 2; Length 1091;
 Best Local Similarity 21.3%; Pred. No. 4.2e-16;
 Matches 205; Conservative 92; Mismatches 326; Indels 340; Gaps 33;
 QY 30 LSESLGTLCPFSKGLL-----FVPPDIDRRVVELRIGGNFIHIGRQDFANMTGLVDIT 82
 Db 188 LRSRLTLRLSKNRITQLPVAKPLP----RLTQLDLNRNRIRLIEGLTFQGLDLSLEVL 243
 QY 83 LSRNTLSHTQPPSFLDESRLSRLHLSNRLPSLCEDTLRLVNLQHLIVNNNOLGGIAD 142
 Db 244 LQRNNISRTDGAFWGLSKMHVJLHLEYNLSVEVNSGLYGLTALHQLHLSNNSISRIQR 303
 QY 143 APEDFLTLEDLDSVNNLHGLPWDSVRRMVLNQLSLDHLNLDHIAEAGTFADLOKLARL 202
 Db 304 GW-SFCCKLHELLLSFNLLRDEESLAEVLSLISLRLSHNAISHAIEAGFKLKSRLV 362
 QY 203 DLTSNRLQKLPDPIFARSOASALT-ATPPAPPLS-----FSFGNPLH- 245

Db 363 DLDHNEISGITEEDTSGAFTGLDNLNLSKLTFLGNKIKSVAKRAFSGLSEHLEHLNGBENAIRS 422
 QY 246 -----CNCCELLML-----RRLERDDDDLETCGSPGGLKGRFYW 277
 Db 423 VQDFAFAMKMLKELYISSESFCLDCQLKWLPPMLMGRMLQAFVTATCAHPESLKGQSI 482
 QY 278 HVRSEEFVCE-----PPLITO----- 293
 Db 483 SVLPDSFVCDFFPKPQIITQPEITMVAVVGKDIRFTCSAASSSSSPMTFAWKDNEVLANA 542
 QY 294 -----HTHK----- 297
 Db 543 DMENFAHVAQDGEVMEYTTILHLRHVTFGHEGRYQCIIINHFGTYSHKARLTVNVLPS 602
 QY 298 -----LLVLEGOAATLKCKAIGDPSPLIHW-----VAPDORLVGNS 333
 Db 603 FTKPHDIAIRGTTRALECAATGHPNPQIAWQKGGTDFPAARERRMHVMPDD----- 656
 QY 334 SRTAVYDNGTLDIIFITTS---QDSGAFTCIAANAAGEATAMVEVSIQVLPKLSNSTRTA 390
 Db 657 -----DVFPIITDKIDDMGVYSCTAQSAGSVSANATLTVLETPSL-----A 698
 QY 391 PPKSRLSDITGSSKTSRGGGGGGGPPKPPERAVLVSEVTTTSALVKSIVKSAPRVK 450
 Db 699 VPLE-----DRVVTVGE---TVAFOCKATGSPTPRIT 727
 QY 451 MYQLQVNCSDDEVLYRMI PASNKAFVNNLVSGTYDVCVLAWDDTAT-TLTATNIVG 509
 Db 728 WLKGRPLSLTE--RHHFTP-GNQLLVVQNV-----MIDDAGRYTCENSNPLG 772
 QY 510 CAQFFTKADY---PQCOSMHSQILGGTMMILVIGIIVATLVFVILMVRVYKVCNHEAPS 566
 Db 773 TERHSQLSILPTFCRCK-DGTTVGI FTIAVVCISVILTSLVWVVCIIYQTRKK----- 823
 QY 567 KMAAAVSNVYSGTNGAQQPPSSAPAGAPPOGP-----PKVVVRNB----- 607
 Db 824 -----SBEYSVNTDETIVPPDVPVSYLSQGLSDRQETVVRTEGGHQANGHIESNGVC 877
 QY 608 LLDFTASLARASDSSSSSLGSGEAGLGRAPWRIPPSAPR-PKPSLDRLMGAF---AS 662
 Db 878 LRD--PSLFPVEVDLHSTTCRQPKLCVGYTRFPWKTEKADRTAAPHHTAHSGSAVCSDCS 935
 QY 663 LDLSQRKEELLDSRTPAGRAGTSARGHSDREPLLGGPPAARARSLLPLPLEGKAKRSH 722
 Db 936 TDTAVHPQVPRDSQGP-GTASSQBELRQHDREYSP-----HHPYSGTADGSH 981
 QY 723 SFDMGDFAAAAGGVVPCGYSPPRKVSNIWTKRSLVSNVGMLLPFEEIDLVGARCTFGSS 782
 Db 982 TL-----SGGSLYPSNHD-----RILPSLKNKAASADGN-GDSS 1014
 QY 783 WVM 785
 Db 1015 WTL 1017
 RESULT 3
 T46266
 hypothetical protein DKF2p761A179.1 - human (fragment)
 C:Species: Homo sapiens (man)
 C:Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 04-Feb-2000
 C:Accession: T46266
 R:Blum, H.; Bauersachs, S.; Mewes, H.W.; Cassenhuber, J.; Wiemann, S.
 submitted to the Protein Sequence Database, January 2000
 A:Reference number: Z23034
 A:Accession: T46266
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-421 <AAA>
 A:Cross-references: EMBL:AL137451
 A:Experimental source: adult amygdala; clone DKF2p761A179
 C:Genetics:
 A:Note: DKF2p761A179.1

QY 295 THKLLVLEGOAATLKCKA 312
 Db 285 -----APMECGA 291

RESULT 6
 A36665
 slit protein 1 precursor - fruit fly (*Drosophila melanogaster*)
 C:Species: *Drosophila melanogaster*
 C:Date: 30-Apr-1991 #sequence_revision 30-Apr-1991 #text_change 02-Aug-2002
 C:Accession: A36665; A31640; S13523
 R:Rothberg, J.M.; Jacobs, J.R.; Goodman, C.S.; Artavanis-Tsakonas, S.
 Genes Dev. 4, 2169-2187, 1990
 A:Title: slit: an extracellular protein necessary for development of midline glia and cell
 A:Reference number: A36665; MUID:91099665; PMID:2176636
 A:Accession: A36665
 A:Status: Preliminary
 A:Molecule type: mRNA
 A:Residues: 1-1480 <ROT>
 A:Cross-references: GB:X53959; NID:98614; PIDN:CRA37910.1; PID:98615
 R:Rothberg, J.M.; Hartley, D.A.; Waither, Z.; Artavanis-Tsakonas, S.
 Cell 55, 1047-1059, 1988
 A:Title: slit: An EGF-homologous locus of *D. melanogaster* involved in the development of
 A:Reference number: A31640; MUID:89077533; PMID:3144436
 A:Accession: A31640
 A:Molecule type: DNA
 A:Residues: 881-1182, 'G', 1185-1404, 'GT', 1463-1464, 'YHA', <RO2>
 A:Cross-references: GB:M23543; NID:G340939; PID:G514357
 C:Genetics:
 A:Gene: FlyBase:slit
 A:Cross-references: FlyBase:FBgn0003425
 A:Introns: 1351/3
 C:Superfamily: fruit fly slit protein; EGF homology; leucine-rich alpha-2-glycoprotein 1

QY 59 LGGNFIIHSRODFANMTGLVDLTLRSNTISHIQPFSLDLESLSLHSDNRLSLGSD 118
 Db 107 LQGNLTVIYETDFQRLTKRLMLQLTDNQIHTIERNISFDLVSLESLERLDSINNVITVGR 166
 QY 119 TLRLGLVNLQHLIYNNLQGGIADAEAFEDLLEDLSDSYNNLHGLPWSVRRMVLHQL 178
 Db 167 VFKAQSLRSLOLDNNQITCLDEHAFKGLVELEILTLLNNNLTSLP----- 212
 QY 179 SLDDHLLDHIAGTFADLQKLARLDTLSNRLQKLPDPPIFARSQASALTATFPAPLSFS 238
 Db 213 ---HNI-----FCGLGRRLRALRSDN-----PFA----- 233
 QY 239 FGGNPLHFCNCELWLRLRERD----DDLETQSGPGKGRYFWHVREBFVCEPPLITQH 294
 Db 234 -----CDCHLSWLSRFLRSATRLAPYTRCQSPQLKQGNVADLHDQFKCSG--LTEH 284
 QY 295 THKLLVLEGOAATLKCKA 312
 Db 285 -----APMECGA 291

RESULT 7
 T13953
 MEGF5 protein - rat
 N:Alternate names: slit protein homolog
 C:Species: *Rattus norvegicus* (Norway rat)
 C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 16-Aug-2002
 C:Accession: T13953
 R:Nakayama, M.; Nakajima, D.; Nagase, T.; Nomura, N.; Seki, N.; Ohara, O.
 Genomics 51, 27-34, 1998
 A:Title: Identification of high-molecular-weight proteins with multiple EGF-like motifs
 A:Reference number: Z14126; MUID:98360089; PMID:9693030
 A:Accession: T13953
 A:Status: Preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-1523 <NNA>
 A:Cross-references: EMBL:AB011531; NID:G3449291; PIDN:BA032461.1; PID:G3449292
 C:Genetics:
 A:Gene: MEGF5
 C:Superfamily: fruit fly slit protein; EGF homology; leucine-rich alpha-2-glycoprotein 1

Query Match 6.3%; Score 259; DB 2; Length 1523;
 Best Local Similarity 27.3%; Pred. No. 9.9e-09;
 Matches 74; Conservative 35; Mismatches 94; Indels 68; Gaps 4;

QY 20 ACPKYCVQNLSESLGTLCPKGLLFPVPPIDRRTVELRGGNFIIHSRODFANMTGLV 79
 Db 279 SCPASCSNNIVD---CRGKGLTEIPANLPEGIIVEIRLEQNSIKSIPAGAFIOYKKLK 334
 QY 80 DLTLSRNTISHIQPFSLDLESLSLHSDNRLSLGSDTLRGLVNLQHLIYNNLQGGI 139
 Db 335 RIDISKNOISDIAPDAFQGLKSLTSLVLYGNKITEIPKGLDFGLVSLQLLLLNANKINCL 394
 QY 140 ADEAFEDFULTLEDLSDSYNNLHGLPWSVRRMVLHQLSDHLLDHIAGTFADLQKL 199
 Db 395 RVNTEQD-----LQNLNLSLYDNKLTQISKGLFAPLQSI 429
 QY 200 ARLDLTSNRLQKLPDPPIFARSQASALTATFPAPLSFSFGGNPLHCCNELLWLRRLERD 259
 Db 430 QTLHLAQ-----CGSPGKGRYFWHVREBFVC 286
 QY 260 DDLET---CGSPGKGRYFWHVREBFVC 286
 Db 455 NPFIETSGARFCSFRRLANKRISQIKSKKFC 485

RESULT 8
 T28714
 hypothetical protein T28714.9a - *Caenorhabditis elegans*
 C:Species: *Caenorhabditis elegans*
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 29-Oct-1999
 C:Accession: T28714
 R:Woessner, J.

QY 6 GGLAFAFMF-----AVVDA-CPKYCVQNLSESLGTLCPKGLLFPVPPIDRRTVELR 58
 Db 51 GGLGSGVGHIPGGVGVIETARCFVCSCTGLNVD----CSHRGLTSVPRKISADVERLE 106

Query Match 6.4%; Score 263; DB 2; Length 1480;
 Best Local Similarity 28.0%; Pred. No. 5.2e-09;
 Matches 89; Conservative 39; Mismatches 102; Indels 88; Gaps 11;

QY 66-91/Domain: proteoglycan amino-terminal homology <PAH1>
 F:101-124/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR11>
 F:125-148/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR12>
 F:149-172/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR13>
 F:173-196/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR14>
 F:197-220/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR15>
 F:221-244/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR16>
 F:245-268/Domain: proteoglycan carboxyl-terminal homology <PCS1>
 F:269-292/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR17>
 F:293-316/Domain: proteoglycan amino-terminal homology <PAH2>
 F:317-340/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR18>
 F:341-364/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR19>
 F:365-388/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR20>
 F:389-412/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR21>
 F:413-436/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR22>
 F:437-460/Domain: proteoglycan carboxyl-terminal homology <PCS2>
 F:461-484/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR23>
 F:485-508/Domain: proteoglycan amino-terminal homology <PAH3>
 F:509-532/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR24>
 F:533-556/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR25>
 F:557-580/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR26>
 F:581-604/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR27>
 F:605-628/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR28>
 F:629-652/Domain: proteoglycan carboxyl-terminal homology <PCS3>
 F:653-676/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR29>
 F:677-700/Domain: proteoglycan amino-terminal homology <PAH4>
 F:701-724/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR30>
 F:725-748/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR31>
 F:749-772/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR32>
 F:773-796/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR33>
 F:797-820/Domain: proteoglycan carboxyl-terminal homology <PCS4>
 F:821-844/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR34>
 F:845-868/Domain: proteoglycan amino-terminal homology <PAH5>
 F:869-892/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR35>
 F:893-916/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR36>
 F:917-940/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR37>
 F:941-964/Domain: proteoglycan carboxyl-terminal homology <PCS5>
 F:965-988/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR38>
 F:989-1012/Domain: proteoglycan amino-terminal homology <PAH6>
 F:1013-1036/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR39>
 F:1037-1060/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR40>
 F:1061-1084/Domain: proteoglycan carboxyl-terminal homology <PCS6>
 F:1085-1108/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR41>
 F:1109-1132/Domain: proteoglycan amino-terminal homology <PAH7>
 F:1133-1156/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR42>

submitted to the EMBL Data Library, August 1997
 A:Description: The sequence of C. elegans cosmid T21D12.
 A:Reference number: Z20514
 A:Accession: T28714
 A:Status: preliminary; translated from GB/EMBL/DBDJ
 A:Molecule type: DNA
 A:Residues: 1-789 <WOE>
 A:Cross-references: EMBL:AF016687; PIDN:AAC48096.1; GSPDB:GN00022; CESP:T21D12.9a
 A:Experimental source: strain Bristol N2; clone T21D12
 C:Genetics:
 A:Gene: CESP:T21D12.9a
 A:Map position: 4
 A:Introns: 38/2; 84/2; 132/2; 204/2; 275/2; 351/2; 519/3; 615/1; 714/1; 758/2

Query Match 6.2%; Score 254; DB 2; Length 789;
 Best Local Similarity 21.3%; Pred. No. 8.5e-09;
 Matches 131; Conservative 75; Mismatches 204; Indels 204; Gaps 25;

QY	56	ELRLCGNFIHSRODFANMTGLVDLTLSRNVTISHIQPFSDLESLSLHLS	-----	109
DB	83	KLDLASNSITDGTDFSSFNLTCLKARHNHTLNQFSRLKLESLLTRNMIREV	142	
QY	110	-----NRPLSGEDTLR-----GLVNIQHLIIVNNQGGIADAFDFLTL	151	
DB	143	RFLAFNQPSLQNVSLARNDVRLDDGMFYACEGLKHLNLSNRVQAVT-EGWMEFLTSL	201	
QY	152	EDDLDSYN-----NLH-----GLPWSVRRMVNHLQSLDHNLDDH	187	
DB	202	EVLDSLQSYQSFHSISSHSHTPKLWLSLHNSRIOSLPSGFRVLRQLEELLSANSIDS	261	
QY	188	TAEGTFADLQKLARLDLTSNRLO-----KLPPDPPIFAR	220	
DB	262	LHKFALVGMSSLHLKDLSSNTLAVCEGAVLYNTSMPLSLRFTNNQLRVIPKRAFER	321	
QY	221	SOASA--LTATPFA-----PPLSFS--FGGNPLHCNCCELLMLR-----	261	
DB	322	FPALDELDTNPDIATIHPEAEPPELEKRLVMNSSSILCQCISWLASWIYRKLKDKSSI	381	
QY	262	LETCSGGLKGRYFWHREBEFVC--EPP--LITQHTKLLVLEGOAATLKCKAIGDPS	317	
DB	382	IAKCSYPPPLADLVVAIDTANLTCNDSRPRAKIVRQPVVESTLIGERKARFTCNVYG--	440	
QY	318	PL-IHW-----VAPDDRVLGNSRSTAVYDNGTLD-----IFITTSODSGA	356	
DB	441	PLSIEWRVMEQPRVLVODSATFLSINRTAVV-NGTFDERELAAAALLLDNVAMTDNSE	499	
QY	357	FTCIAAANAAG-EATAMVEVSIVQLPHLSNSTRTAPPKRSRLSDITGSSKTSRGGGGGG	415	
DB	500	YOCVARNRFGSDFSTHVKLVYVQAPKFT-----	527	
QY	416	EPPKSPPERAVLVSEVT-----TTSALVKWSYK-SAPRVKMYQIYNGCSDDDEVLIY	466	
DB	528	---YTPEMDPLLVGQTAKFLCAATGTPPEIKWAEQIPFPFAEARRLVVTPDDHI---	581	
QY	467	RMIPASNAKAFVNNLVSGTGYDLCLVAMWDDTATLTATNIVGCAQ-----	516	
DB	582	---YIMN-----VTKEDQAYTCHATNAGQTAQASANLVIVFENFF--	618	
QY	517	ADYPOCQSMHSQIL	530	
DB	619	-HYPEPDLSPMLI	631	

RESULT 9

T28715
 hypothetical protein T21D12.9b - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 29-Oct-1999
 C:Accession: T28715
 R:Woessner, J.
 submitted to the EMBL Data Library, August 1997
 A:Description: The sequence of C. elegans cosmid T21D12.
 A:Reference number: Z20514

A:Accession: T28715
 A:Status: preliminary; translated from GB/EMBL/DBDJ
 A:Molecule type: DNA
 A:Residues: 1-1355 <WOE>
 A:Cross-references: EMBL:AF016687; PIDN:AAC48095.1; GSPDB:GN00022; CESP:T21D12.9b
 A:Experimental source: strain Bristol N2; clone T21D12
 C:Genetics:
 A:Gene: CESP:T21D12.9b
 A:Map position: 4
 A:Introns: 38/2; 84/2; 132/2; 204/2; 275/2; 351/2; 519/3; 615/1; 714/1; 758/2; 786/3; 84

Query Match 6.2%; Score 254; DB 2; Length 1355;
 Best Local Similarity 21.3%; Pred. No. 1.8e-08;
 Matches 131; Conservative 75; Mismatches 204; Indels 204; Gaps 25;

QY	56	ELRLCGNFIHSRODFANMTGLVDLTLSRNVTISHIQPFSDLESLSLHLS	-----	109
DB	83	KLDLASNSITDGTDFSSFNLTCLKARHNHTLNQFSRLKLESLLTRNMIREV	142	
QY	110	-----NRPLSGEDTLR-----GLVNIQHLIIVNNQGGIADAFDFLTL	151	
DB	143	RFLAFNQPSLQNVSLARNDVRLDDGMFYACEGLKHLNLSNRVQAVT-EGWMEFLTSL	201	
QY	152	EDDLDSYN-----NLH-----GLPWSVRRMVNHLQSLDHNLDDH	187	
DB	202	EVLDSLQSYQSFHSISSHSHTPKLWLSLHNSRIOSLPSGFRVLRQLEELLSANSIDS	261	
QY	188	TAEGTFADLQKLARLDLTSNRLO-----KLPPDPPIFAR	220	
DB	262	LHKFALVGMSSLHLKDLSSNTLAVCEGAVLYNTSMPLSLRFTNNQLRVIPKRAFER	321	
QY	221	SOASA--LTATPFA-----PPLSFS--FGGNPLHCNCCELLMLR-----	261	
DB	322	FPALDELDTNPDIATIHPEAEPPELEKRLVMNSSSILCQCISWLASWIYRKLKDKSSI	381	
QY	262	LETCSGGLKGRYFWHREBEFVC--EPP--LITQHTKLLVLEGOAATLKCKAIGDPS	317	
DB	382	IAKCSYPPPLADLVVAIDTANLTCNDSRPRAKIVRQPVVESTLIGERKARFTCNVYG--	440	
QY	318	PL-IHW-----VAPDDRVLGNSRSTAVYDNGTLD-----IFITTSODSGA	356	
DB	441	PLSIEWRVMEQPRVLVODSATFLSINRTAVV-NGTFDERELAAAALLLDNVAMTDNSE	499	
QY	357	FTCIAAANAAG-EATAMVEVSIVQLPHLSNSTRTAPPKRSRLSDITGSSKTSRGGGGGG	415	
DB	500	YOCVARNRFGSDFSTHVKLVYVQAPKFT-----	527	
QY	416	EPPKSPPERAVLVSEVT-----TTSALVKWSYK-SAPRVKMYQIYNGCSDDDEVLIY	466	
DB	528	---YTPEMDPLLVGQTAKFLCAATGTPPEIKWAEQIPFPFAEARRLVVTPDDHI---	581	
QY	467	RMIPASNAKAFVNNLVSGTGYDLCLVAMWDDTATLTATNIVGCAQ-----	516	
DB	582	---YIMN-----VTKEDQAYTCHATNAGQTAQASANLVIVFENFF--	618	
QY	517	ADYPOCQSMHSQIL	530	
DB	619	-HYPEPDLSPMLI	631	

RESULT 10

JC1282
 insulin-like growth factor-binding protein acid labile chain precursor - rat
 C:Species: Rattus norvegicus (Norway rat)
 C>Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 21-Jul-2000
 C:Accession: JC1282
 R: Dai, J.; Baxter, R.C.
 Biochem. Biophys. Res. Commun. 188, 304-309, 1992
 A:Title: Molecular cloning of the acid-labile subunit of the rat insulin-like growth fac
 A:Reference number: JC1282; MUID:93038676; PMID:1384485
 A:Accession: JC1282
 A:Molecule type: mRNA
 A:Residues: 1-603 <DAI>

A;Cross-references: GB:S46785; NID:g258002; PIDN:AAB23770.2; PID:g5705934
 A;Experimental source: liver
 A;Note: the authors translated the codon AAG for residue 63 as Arg, AAA for residue 205
 C;Superfamily: leucine-rich alpha-2-glycoprotein repeat homology
 F:1-27/Domain: signal sequence #status predicted <SIG>
 F:28-603/Product: insulin-like growth factor binding protein, acid labile chain #status
 F:267-290/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR>

Query Match 6.1%; Score 248; DB 2; Length 603;
 Best Local Similarity 25.2%; Pred. No. 1.5e-08;
 Matches 104; Conservative 64; Mismatches 161; Indels 84; Gaps 15;

QY 8 LLAFGMAF-----AVVDA---CPKYVCV--QNLSSEGLTCLPCKGLFVPPDI 50
 DB 13 LLAFWALGCHLQGTDFGASADAEQPCPVACTSHDDYDDELVSFCCSKNLTLPDDI 72
 QY 51 DRRTVELRGGNFIIHISRODFANMTGLVDLTLRSNTTSHIQFFSFLDLESURSLLHDSN 110
 DB 73 PVSTRALWLDGNNLSSIPAAQNLSSLDFFLNQGSWLRSLRLEPQALLGLQNLIIYHLERN 132
 QY 111 RLPSLGEDTLRGLVNLQHLIYVNNQGGIADAEFDFLTLDELDSYNNLHGLPWSVR 170
 DB 133 LRNLAVGLFTTPSLASLSLSLNLGRLEEGFQG-LSHLWDLNLGWNLSVLLPDTVFQ 191
 QY 171 RMVNLHOLSLDHNLHDIAEAGTFADLQKLRDLTSNRLOKLPDPPIFARSOASALTATP 230
 DB 192 GLGNLHELVLGNKLYLQPALFCGLGELRELDLSRNLRSV-----KANVFLHP 242
 QY 231 FAPPLSFGGNPLHNCBELLMLRERODDLETGCGSPGLKG--RYFW----HVR---- 280
 DB 243 -----RLQKLYLDRNLITAVAGFLGKALRWLDLSHNRVAGL 281
 QY 281 -EEFVCEPPLITQHTHKLVLBQQAATLCKRAIGDPPSLIHWVAPDRLVGNSSRTAVY 339
 DB 282 MEDTF---PCLGLGHV--LRLAHNAIASLRPTFKD----LHPL--EELQGHNRIRQLG 330
 QY 340 DN-----GLDIF-----ITTSQSGAFTCIAANAAGATAMVEVSIQV 378
 DB 331 ERTFEGIGQLVLTLDNQITEVRVGFAGSGLFNVAVMNLGNCNLRSLRPERVFG 383

RESULT 11
 A41915
 insulin-like growth factor-binding complex acid-labile chain precursor - human
 N;Alternate names: Acid-Labile Subunit (ALS)
 C;Species: Homo sapiens (man)
 C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 05-Nov-1999
 C;Accession: A41915
 R;Leong, S.R.; Baxter, R.C.; Camerato, T.; Dai, J.; Wood, W.I.
 Mol. Endocrinol. 6, 870-876, 1992
 A;Title: Structure and functional expression of the acid-labile subunit of the insulin-1
 A;Reference number: A41915; MUID:92357025; PMID:1379671
 A;Accession: A41915
 A;Status: preliminary
 A;Molecule type: mRNA; protein
 A;Residues: 1-605 <LEO>
 A;Cross-references: GB:M86826; NID:g184807; PIDN:AAA36047.1; PID:g184808
 A;Experimental source: liver
 A;Note: sequence extracted from NCEI backbone (NCBI:P110171)

C;Superfamily: leucine-rich alpha-2-glycoprotein repeat homology
 F;75-98/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR1>
 F;99-122/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR2>
 F;123-146/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR3>
 F;147-170/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR4>
 F;171-194/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR5>
 F;195-218/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR6>
 F;219-242/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR7>
 F;243-266/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR8>
 F;267-290/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR9>
 F;291-314/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR10>
 F;315-338/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR11>
 F;339-362/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR12>
 F;363-386/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR13>

F;387-410/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR14>
 F;411-434/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR15>
 F;435-458/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR16>
 F;459-482/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR17>
 F;483-506/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR18>
 F;507-529/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR19>

Query Match 5.9%; Score 241.5; DB 2; Length 605;
 Best Local Similarity 29.7%; Pred. No. 3.9e-08;
 Matches 76; Conservative 41; Mismatches 106; Indels 33; Gaps 4;

QY 20 ACPKYVCV--QNLSSEGLTCLPCKGLFVPPDIDRRVVELRGLGNFIHISRODFANMTG 77
 DB 40 ACPAAVCVSDDDADELSVFCSSRNRLTRPDGPPGQTOALWLDGNNLSSVPPAAFOQLSS 99
 QY 78 LVDLTLRSNTTSHIQFFSFLDLESRLSLHDSNRLPSLGEDTLRGLVNLQHLIYVNNQGL 137
 DB 100 LGFNLQGGQLSLEPQALLGLENLCHLHLERNQLRSALGTFAHTPALASLGLSNRRLS 159
 QY 138 GIADAEFDFLTLDELDSYNNLHGLPWSVRMNLHOLSLDHNLHDIHAEGTFFADLQ 197
 DB 160 RLEDGLFEG-LGSLWDLNLGWNLSLAVLPDAAFRGLGSLRELIVLAGNRLAVLQPALFSGLA 218
 QY 198 KLARLDLTSNRLOKLPDPPIFARSOASALTATPAPPLSFGGNPLHNCBELLWLRLE 257
 DB 219 ELRELDLSRNALR-----AIKANVFEV-----OLPRLQKLY 248
 QY 258 RDDDELETGCGSPGLKG 273
 DB 249 LDRNLIAVAVGAFGLG 264

RESULT 12
 T42626
 secreted leucine-rich repeat-containing protein SLIT2 - mouse (fragment)
 N;Alternate names: neurogenic extracellular slit protein
 C;Species: Mus musculus (house mouse)
 C;Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 16-Aug-2002
 C;Accession: T42626
 R;Holmes, G.P.; Negus, K.; Burridge, L.; Raman, S.; Algar, E.; Yamada, T.; Little, M.H.
 Mech. Dev. 79, 57-72, 1998
 A;Title: Distinct but overlapping expression patterns of two vertebrate slit homologs in
 A;Reference number: 222177; MUID:99279238; PMID:10349621
 A;Accession: T42626
 A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: mRNA
 A;Residues: 1-1025 <HOL>
 A;Cross-references: EMBL:AF074960; NID:g4151258; PID:g4151259; PIDN:AAD04345.1
 C;Genetics:
 A;Gene: Slit2
 C;Superfamily: fruit fly slit protein; EGF homology; leucine-rich alpha-2-glycoprotein r

Query Match 5.8%; Score 237; DB 2; Length 1025;
 Best Local Similarity 20.4%; Pred. No. 1.6e-07;
 Matches 86; Conservative 60; Mismatches 110; Indels 166; Gaps 10;

QY 20 ACPKYVCVQNLSSEGLTCLPCKGLFVPPDIDRRVVELRL----- 59
 DB 1 ACEPKRCRCE-----GTTVDCSNQRLNKPDPHPOYTAELRLNNEFTVLEATGIFKXLP 54
 QY 60 -----CGNFIHISRODFANMTGLVLTLSRNTTSHIQFFSFLDLESRLSLHDSNRL 112
 DB 55 QLRXINFNKNIITDEEAFEGAGSAGVNEILLTSNRLENVQHKMFKGLSESLKTLMLRSNRI 114
 QY 113 PSLGEDTLRGLVNLQHLIYVNNQGGIADAEFDFLTLDELDSYNNLH-----NLH----- 162
 DB 115 SCVGNDSFTGLGSRVLLSYDNIITVAPGAF-DXLHLSLTLNLLANPFCNCHLAWLGE 173
 QY 163 ----- 162
 DB 174 WLRRKRIVTGNPRCKQPKYFLKEIPQDVAIQDFTCDGNDNSCSPLRSPSECTCLDTX 233
 QY 163 -----GLPWSVRRMNLHQ-----LSLDHNLHDIAEAGTF 193

Db 234 VRCSNKGKLVLPKGIKPKVDTELYDGNQFTLVKPKELSNYKHHTLIDLGNRRISTLSNQXF 293
 QY 194 ADLQKLRDLTNSRLOKLPDP-----PIFARSOAGALTATPPA 232
 Db 294 SNMTQLLTLLSYNRLRCPPTFDGLKSLRLLSLHGDINDISVVPEGAFNDLSALS----- 348
 QY 233 PPLSFGNPHHCNCELLWL-----RRLERDDLETGCGSPGGLKGRYFWMHREBEFVCEP 288
 Db 349 ---HLAIGANPLCYDCNMOWLSDDWVKSEYKEPGIARCAAGPGEMADKLLITTPSKKFTCOG 405
 QY 289 PL 290
 Db 406 PM 407

RESULT 13
 AS3860
 chondroaderin precursor - bovine
 N:Alternate names: 38K leucine-rich protein
 C:Species: Bos primigenius taurus (cattle)
 C:Date: 07-Oct-1994 #sequence_revision 07-Oct-1994 #text_change 05-Nov-1999
 C:Accession: A53860
 R:Name: P.J.; Sommarin, Y.; Boynton, R.E.; Heinegard, D.
 J. Biol. Chem. 269, 21547-21554, 1994
 A:Title: The structure of a 38-kDa leucine-rich protein (chondroaderin) isolated from b
 A:Reference number: A53860; MUID:94342341; PMID:8063792
 A:Accession: A53860
 A:Status: Preliminary
 A:Molecule type: mRNA
 A:Residues: 1-361 <NEA>
 A:Cross-references: GB:U08018; MID:9470671; PIDN:AAA21330.1; PID:9470672
 C:Superfamily: leucine-rich alpha-2-glycoprotein repeat homology; proteoglycan carboxyl-
 C:Keywords: disulfide bond
 F:300-346/Domain: proteoglycan carboxyl-terminal homology <PCH>

Query Match 5.8%; Score 236; DB 2; Length 361;
 Best Local Similarity 25.4%; Pred. No. 4.4e-08;
 Matches 90; Conservative 42; Mismatches 122; Indels 100; Gaps 11;

QY 7 GLLAFGMAFVVDACPKYCVQNLSESLGTLCPKGLLFPVPPDIDRRVLELGGNFIIH 66
 Db 13 GLLA--SLLPALAACPQCHCH--SDLQHVICDKVGLQKI-PKVBSEKTLMLQRNPFV 67
 QY 67 ISRODFANMTGLVD-----LTLRSNTISHIQPFSFLDLSL 102
 Db 68 LATNSFRAMPNLVSLHLQHCQIREVAAGAFRGLKQLIYLYLSHNDIRVLRAGAFDLDLTEL 127
 QY 103 RSLHLDNSRPLSGEDTTLRGLVNLQHLIYVNNQGLGIADEAFE----- 145
 Db 128 TVLYLDHNKVTLEPRGLLSPVNLFTLQNNKIRELRSGAFQGAQKDLRWLYLSENSLSS 187
 QY 146 -----DELLLEDLDSYNNLHGLPWSVRRMWN- 174
 Db 188 LQFGALDDVENLAKFYLDNRNQLSSVPSAALSRLRVVEELKLSHPKSIIPDNAFOSGRY 247
 QY 175 LHLQSLDHLNLDHIAEGTADLQKLRDLTNSRLOKLPDPPIFARSOASALTATPPAPP 234
 Db 248 LETLWLDNTNLEKFSGDGALGVTTLKHVILENNRNLHOLPSN--FPDLSLETILT----- 300

QY 235 LSFSGGNPLHCNCELLMWR-----LERDDLETGCGSPGGLKGRYFWMHREBE 283
 Db 301 -----NNPWKCTCOLRGLRRLEAKTSRPD--ATCASPAKFRGQ---HIROTD 343

RESULT 14
 JC5239
 insulin-like growth factor acid-labile chain - baboon
 C:Species: Papio sp. (baboon)
 C:Date: 17-Apr-1997 #sequence_revision 09-May-1997 #text_change 09-May-1997
 C:Accession: JC5239
 R:Delhanty, P.; Baxter, R.C.
 Biochem. Biophys. Res. Commun. 227, 897-902, 1996

A:Title: The cloning and expression of the baboon acid-labile subunit of the insulin-like
 A:Reference number: JC5239; MUID:97040714; PMID:8886027
 A:Contents: liver
 A:Accession: JC5239
 A:Molecule type: mRNA
 A:Residues: 1-605
 C:Comment: This factor is structurally related to proinsulin and have insuline-like meta

Query Match 5.8%; Score 235.5; DB 2; Length 605;
 Best Local Similarity 30.5%; Pred. No. 9.8e-08;
 Matches 78; Conservative 34; Mismatches 111; Indels 33; Gaps 4;

QY 20 ACPKYCVQNLSE--SLGTLCPKGLLFPVPPDIDRRVLELGGNFIIHISRODFANMTG 77
 Db 40 ACPATCACSYDDEVNELSVFCSSRNLTLPDIPGIGTQALWLDSDNSNLSIPPAFRNLSS 99
 QY 78 LVDTLTSRNTISHIQPFSFLDLSRSLHLDNSRPLSGEDTLRGLVNLQHLIYVNNQGL 137
 Db 100 LAFLNLQGGQGSLEPQALLGLENLCHUHLERNQRSLAVGTFATPPALALLGLSNNRUS 159
 QY 138 GIADAEAFEDFLTTLEDLDSYNNLHGLPWSVRRMWNLHQLSLDHLNLDHIAEGTFADIQ 197
 Db 160 RLEDGLFEG-LGNLWDLNLGNWNLAVLPDAARFGLGGLRELVLGNRLAYLQPALFSGLA 218
 QY 198 KLARLDLTSRNLQKLPDPIFARSOASALTATPPAPPISFSGGNPLHCNCELLMWRLE 257
 Db 219 ELRELDLSRNALR-----AIKANVFA-----OLPRLQKLY 248
 QY 258 RDDLETGCGSPGGLKG 273
 Db 249 LDRNLIATAAVPAGAFLG 264

RESULT 15
 S46224
 peroxidase - fruit fly (Drosophila sp.)
 C:Species: Drosophila sp.
 C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 24-Oct-2000
 C:Accession: S46224
 R:Nelson, R.E.; Fessler, L.I.; Takagi, Y.; Blumberg, B.; Keene, D.R.; Olson, P.F.; Parke
 EMBO J. 13, 3438-3447, 1994
 A:Title: Peroxidase: a novel enzyme-matrix protein of Drosophila development.
 A:Reference number: S46224; MUID:94341255; PMID:8062820
 A:Accession: S46224
 A:Status: Preliminary
 A:Molecule type: mRNA
 A:Residues: 1-1535 <NEL>
 C:Superfamily: peroxidase; myeloperoxidase homology; proteoglycan amino-terminal homolo
 F:19-44/Domain: proteoglycan amino-terminal homology <PAH4>
 F:661-1350/Domain: myeloperoxidase homology <MPX>

Query Match 5.7%; Score 232.5; DB 2; Length 1535;
 Best Local Similarity 23.5%; Pred. No. 5.5e-07;
 Matches 102; Conservative 56; Mismatches 193; Indels 83; Gaps 13;

QY 1 METLGGLLAFGMAFVVDACPKYCVQNLSESLGTLCPKGLLFPV--PDIDRRVLEUR 58
 Db 7.MQLLGLLALLAGGVQSV-YCPAGCTCLERT----VRCIRAKLSAVKLPQ-DTQFLDLR 60

QY 59 IGGNFIIHISRODFANMTGLVDLTSRNTISHIQPFSFLDLSRSLHLDNSRPLSGED 118
 Db 61 F--NHIEELPANAFSGLAQLTTLFLNDNELAYLQDGLNGLTALRFVYVNNRNLRLPAT 118

QY 119 TLRGLVNLQHLIYVNNQGLGIADEAFEDFLTTLEDLDSYNNLHGLPWSVRRMWNLHQL 178
 Db 119 IFQRPRLREGIFLENNDIWQLPAGLFL-DNLPRNLRLIMYNNKLTQPLVDFGNRNLNRL 177

QY 179 SLDHLNLDHIAEGTADLQKLRDLTNSRLOKLPDPIFARSOASALTATPPAPPISFS 238
 Db 178 RLDGNAID----- 185

QY 239 FGGNPLHCNCEL--LWLRRLERDDLE-----TCGSPGGLKGRYFWMHREBEFVCEPPL 290

```

Db      186 .-----IDCNCCVYSLW-RRWHLDVQRQVLVSTLJCAAPQMLQNGQFSSLGEHHFKCAKPO 239
QY      291 ITQHTHKLKLYLEGOAATLKCKAIGDPSPLIHWVAPDDRL-VGNSSRTAVYDNGTLDIFIT 349
Db      240 FLVAPQDAQVAAGQVELSCEVTGLHRPQITWMHNTQELGLEEQCAEILPSSGLLHRSA 299
QY      350 TSODSGAFTCIAANAAGEATAMVEVSIVQLPHLSNSTRTAPPKSRSLDITGSSKT-SRG 408
Db      300 DTSDMGYQCIARNEMGALRSQP-----VRLVWGGNHPLDSPIDARSNQVWADAGTPTHG 355
QY      409 GGGSGGGEPPKSP 422
Db      356 ATPLPSPSPSHSP 369
  
```

Search completed: June 8, 2003, 21:31:21
 Job time : 50 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 8, 2003, 19:06:53 ; Search time 36 Seconds
(without alignments)
909.023 Million cell updates/sec

Title: US-09-831-846-2
Perfect score: 4094
Sequence: 1 METLLGGLLAFMAFAVVDA.....DLVGARGTFGSSEWMESTV 789

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	388.5	9.5	713	GAC1_HUMAN	O75325 homo sapien
2	282.5	6.9	646	FLR1_HUMAN	Q9hzul homo sapien
3	263	6.4	1480	SLLT_DROME	P24014 drosophila
4	248	6.0	359	ALS_RAT	P35859 rattus norv
5	244	6.0	359	CHAD_HUMAN	O15335 homo sapien
6	242.5	5.9	966	Y918_HUMAN	O94991 homo sapien
7	241.5	5.9	605	ALS_HUMAN	P35858 homo sapien
8	240.5	5.9	649	FLR3_HUMAN	Q9nzu0 homo sapien
9	240.5	5.9	951	LGR4_RAT	Q9z2h4 rattus norv
10	238	5.8	660	FLR2_HUMAN	O43155 homo sapien
11	236	5.8	361	CHAD_BOVIN	Q27972 bos taurus
12	236	5.8	977	Y848_HUMAN	O94933 homo sapien
13	235.5	5.8	605	ALS_PAPHA	O02833 papio hamad
14	234.5	5.7	951	LGR4_HUMAN	Q9bxb1 homo sapien
15	232.5	5.7	757	LGR7_HUMAN	Q9hbx9 homo sapien
16	231	5.6	603	ALS_MOUSE	P70389 mus musculus
17	226	5.5	358	CHAD_MOUSE	O55226 mus musculus
18	225	5.5	358	CHAD_RAT	O70210 rattus norv
19	224	5.5	373	ASPN_MOUSE	Q95mq4 mus musculus
20	221	5.4	907	LGR5_HUMAN	O75473 homo sapien
21	220	5.4	379	ASPN_HUMAN	Q9bxx1 homo sapien
22	211	5.2	368	PGS1_XENLA	Q91b75 xenopus lae
23	211	5.2	567	GPV_RAT	O08770 rattus norv
24	208	5.1	907	LGR5_MOUSE	Q921p4 mus musculus
25	208	5.1	2012	DSCA_HUMAN	O60469 homo sapien
26	203	5.0	821	TRKB_MOUSE	P15209 mus musculus
27	201	4.9	626	GPBA_HUMAN	O73359 homo sapien
28	201	4.9	905	TLR3_MOUSE	Q99mb1 mus musculus
29	200.5	4.9	567	GPV_MOUSE	O08742 mus musculus
30	198	4.8	536	CBP8_HUMAN	P22792 homo sapien
31	197	4.8	560	GPV_HUMAN	P40197 homo sapien
32	196	4.8	821	TRKB_RAT	Q63604 rattus norv
33	195	4.8	360	PGS2_BOVIN	P21793 bos taurus

ID	GAC1_HUMAN	STANDARD;	PRT;	713 AA.
34	195	4.8	1097	1 TOLL_DROME
35	194.5	4.8	360	1 PGS2_PIG
36	194.5	4.8	1041	1 TLR8_HUMAN
37	194	4.7	360	1 PGS2_SHEEP
38	191.5	4.7	360	1 PGS2_HORSE
39	190	4.6	737	1 LGR8_MOUSE
40	189.5	4.6	378	1 LUM_MOUSE
41	186.5	4.6	375	1 FMOD_BOVIN
42	186.5	4.6	1315	1 CHAO_DROME
43	186	4.5	360	1 PGS2_CANFA
44	185.5	4.5	354	1 PGS2_MOUSE
45	185.5	4.5	369	1 PGS1_BOVIN

ALIGNMENTS

RESULT 1
GAC1_HUMAN
AC O75325;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Glioma amplified on chromosome 1 protein precursor.
GN GAC1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
[1]
RP SEQUENCE FROM N.A.
RC TISSUE=Glial tumor;
RX MEDLINE=98324709; PubMed=9662332;
RA Malfoy B., Almeida A., Zhu X.X., Vogt N., Tvagi R., Muleris M.,
RA Dutrillaux A.-M., Dutrillaux B., Ross D., Hanash S.;
RT "GAC1, a new member of the leucine-rich repeat superfamily on
RT chromosome band 1q32.1, is amplified and overexpressed in malignant
RT gliomas."
RL Oncogene 16:2997-3002(1998).
CC -!- SUBCELLULAR LOCATION: Type I membrane protein (Potential).
CC -!- TISSUE SPECIFICITY: OVERAMPLIFIED IN MALIGNANT GLIOMAS.
CC -!- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY.
CC -!- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAIN.
CC -!- SIMILARITY: CONTAINS 11 LEUCINE-RICH REPEATS (LRR).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
EMBL; AF030435; AAC39792.1; -;
MIM; 605492; -;
DR InterPro; IPR003006; IG_MHC.
DR InterPro; IPR003598; IG_C2.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR000483; LRR_Cterm.
DR InterPro; IPR000372; LRR_Nterm.
DR InterPro; IPR003592; LRR_out.
DR InterPro; IPR003591; LRR_typ.
DR Pfam; PF000047; Ig_1.
DR Pfam; PF00560; LRR; 9.
DR Pfam; PF01462; LRRNT; 1.
DR Pfam; PF01463; LRRCT; 1.
DR PRINTS; PR00019; LEURICHRPT.
DR SMART; SM00408; IGC2; 1.
DR SMART; SM00370; LRR; 6.
DR SMART; SM00082; LRRCT; 1.
DR SMART; SM00013; LRRNT; 1.
DR SMART; SM00369; LRR_TYP; 2.

KW Immunoglobulin domain; Transmembrane; Glycoprotein; Repeat;
 KW Leucine-rich repeat; Signal;
 FT SIGNAL 1 18
 FT CHAIN 19 713
 FT DOMAIN 19 630
 FT TRANSMEM 631 651
 FT DOMAIN 652 713
 FT REPEAT 92 115
 FT REPEAT 116 139
 FT REPEAT 140 163
 FT REPEAT 165 187
 FT REPEAT 188 211
 FT REPEAT 213 235
 FT REPEAT 236 259
 FT REPEAT 261 283
 FT REPEAT 309 333
 FT REPEAT 334 357
 FT REPEAT 359 385
 FT DOMAIN 438 504
 FT DISULFID 445 497
 FT CARBOHYD 94 94
 FT CARBOHYD 381 381
 FT CARBOHYD 555 555
 FT CARBOHYD 583 583
 SQ SEQUENCE 713 AA; 78798 MW; EC8BC0FD240C9396 CRC64;

Query Match 9.5%; Score 388.5; DB 1; Length 713;
 Best Local Similarity 22.5%; Pred. No. 1e-17;
 Matches 170; Conservative 91; Mismatches 242; Indels 251; Gaps 27;

QY 21 CPKYVCQ-----NLSESLGTLCPKSGKGLFVPPDIDRRTVRLRGGNFIIHISQD 71
 DB 29 CPQACQIRPWYTRSSVREATTVDCNDLFLTAVPPALPAGTQTLLOSNSIVRVDQSE 88

QY 72 FANMTGLVDLTSRNTISHIQPFSLDLESLSLHDSNRLPSLGEDTLRGLVNIQHILV 131
 DB 89 LGYLANLTELDSQNSFSDARDCEHALPQLLSLHEENQLTRLEDHSPFAGLASLQELYL 148

QY 132 NNOLGGIAEAF-----DLQKLARLDLTSN-RLQKLP 214
 DB 149 NHNQLYRIAPRAPGSLNLLRHLNSLRLAIDSRWFEMLPNLEILMGGNKVDAILDMN 208

QY 145 -----EDFLL-----TLEDLDSYNNRHLGHPWDSVRRVMVNLHQLSL 180
 DB 209 FRPLANRLSLVLAGMNLREISDYALBGLQSLSPYDQNLARVPRALAEQVPLKFLDL 268

QY 181 DHNLLDHIABGTF-----DLQKLARLDLTSN-RLQKLP 214
 DB 269 NKNPLQRVGPGDFANMLHLKELGLANNMEEVLSIDKFPALVNLPELTKLDITNPRLSFIHP 328

QY 215 DPFARSOASALTATPFA-----PPL-SFSGGNPLHCCELLW-----LR 254
 DB 329 RAFHLLPQMETLMLNNSALSHQQTVESLPLNQLVGLHGNPDRCDVIRWANATGRVTR 388

QY 255 RLERDDDLTCGSPGKGRYFHWREBF-----VCEPPLITQHTH--KLVLVLEQQAT 307
 DB 389 FIEPQSTL--CAEPPDLQR--LPVREVPFREMTHDC-LPLISPRSPFSLQVSGESMV 442

QY 308 LKCKAIGDPSPLIHWVAPDD-RLVGNSS--RTAVYDNGILDIFITTSQDSGAFTCIAANA 364
 DB 443 LHCRALAEPEIYVYTPAGLRITPAHAGRCRVPYEPGTELELRVTAEBAGLYTCAQNL 502

QY 365 AGEATAMVEYSI-----VOLPHLSN--STSRAPPKRLSDITGSSKT 405
 DB 503 VGHADTKTVVVGALLQPRDEGQGLELRVQETHPHILLSWVTPNTVSNLWSSAS 562

QY 406 SRGGGSGGGEPPKPPERAVLSEVTTTSALVQKWSVKSAPRVKMYQLQYNCSDDEVLI 465
 DB 563 SLRQGA-----TALAR----- 574

QY 466 YRMIPASNAKAVVNLVSGTYDLCVLAWDDDTATLITATNIVGCAQFETKADYPOCQSM 525
 DB 575 ---LPRGTHSYNITRLLQATEFYWACLQVAFADAHTQL-----ACVWARTK-----EATSC 621

QY 526 HSQILGGTMLVIGGIIVATLLVFIILMVRYKVCNHEAFSKMAAAVSNVYSQTNGAOPP 585
 DB 622 H-RALGDRPGLI-----AILLAVLLAAGLAHLGTGQPRKGV-----GGRPL 665

QY 586 PPS-----SAPA-----GAP-----PQPPPKVVVRN 606
 DB 666 PPAWAFWGSAPSVRVVVSAPLVLVLPWNPGRKLP 699

RESULT 2
 FLR1 HUMAN STANDARD; PRT; 646 AA.
 AC OGNZUI;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Leucine-rich repeat transmembrane protein FLR1 precursor
 DE (Fibronectin-like domain-containing leucine-rich transmembrane protein
 FLR1).
 GN Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.; TISSUE SPECIFICITY, AND GYLCOSYLATION.
 RX MEDLINE=20112755; Pubmed=10644439;
 RA Lacy S.E., Bonnemah C.G., Buzney E.A., Kunkel L.M.;
 RT "Identification of FLR1, FLR2, and FLR3: a novel family of
 RT transmembrane leucine-rich repeat proteins.";
 RL Genomics 62:417-426 (1999).
 CC !- FUNCTION: May have a function in cell adhesion and/or receptor
 CC signaling.
 CC !- SUBCELLULAR LOCATION: Type I membrane protein (Probable).
 CC !- TISSUE SPECIFICITY: Expressed in kidney and brain.
 CC !- PTM: N-glycosylated.
 CC !- SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE III-LIKE DOMAIN.
 CC !- SIMILARITY: CONTAINS 10 LEUCINE-RICH REPEATS (LRR).
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.iesb-sib.ch/announce/
 CC or send an email to license@sib-sib.ch).
 CC
 DR EMBL; AF169675; AAF28459.1; ALT_INIT.
 DR Genew; HGNC:3760; FLRT1.
 DR MIM; 604806;
 DR InterPro; IPR003961; FN III.
 DR InterPro; IPR001611; LRR.
 DR InterPro; IPR000483; LRR_Cterm.
 DR InterPro; IPR000372; LRR_Nterm.
 DR InterPro; IPR003591; LRR_out.
 DR Pfam; PF00041; fn3; 1.
 DR Pfam; PF00560; LRR; 7.
 DR Pfam; PF01462; LRRNT; 1.
 DR Pfam; PF01463; LRRCT; 1.
 DR PRINTS; PR00019; LEURICHRPT.
 DR SMART; SM00370; LRR; 1.
 DR SMART; SM00082; LRRCT; 1.
 DR SMART; SM00013; LRRNT; 1.
 DR SMART; SM00369; LRR_TYP; 1.
 DR Transmembrane; Leucine-rich repeat; Repeat; Glycoprotein; Signal.
 FT SIGNAL 1 20
 FT CHAIN 21 646
 FT DOMAIN 21 524
 FT TRANSMEM 525 545
 FT DOMAIN 546 646

FT REPEAT 52 77 LRR 1.
 FT REPEAT 78 98 LRR 2.
 FT REPEAT 99 121 LRR 3.
 FT REPEAT 123 147 LRR 4.
 FT REPEAT 148 169 LRR 5.
 FT REPEAT 170 192 LRR 6.
 FT REPEAT 194 218 LRR 7.
 FT REPEAT 219 241 LRR 8.
 FT REPEAT 242 264 LRR 9.
 FT REPEAT 265 288 LRR 10.
 FT DOMAIN 407 485 FIBRONECTIN TYPE-III. (POTENTIAL)
 FT CARBOHYD 221 221 N-LINKED (GLCNAC. . .) (POTENTIAL)
 FT CARBOHYD 277 277 N-LINKED (GLCNAC. . .) (POTENTIAL)
 SQ SEQUENCE 646 AA; FT2BF5DC3A1392 CRC64;
 Query Match 6.9%; Score 282.5; DB 1; Length 646;
 Best Local Similarity 22.2%; Pred. No. 6.4e-11;
 Matches 145; Conservative 84; Mismatches 218; Indels 205; Gaps 27;
 Oy 7 GLLAFGMAFVDA--CPKYCVONLSLSTGL-CPKSKGLLFVPPDI--DRRTV----- 55
 Db 12 GLIAF--LTEVIDSTCPVCRCDN-----GFIYCNDRGLTSIPADIPDDATTLYLQNNQ 64
 Oy 56 -----ELRGGNFIIHISRODFANWTG 77
 Db 65 INNAGIPQDLKTKVNVQVIYLYENDLDEFFINLRSLREHLQDNVNTIARDSLARIPL 124
 Oy 78 LVDTLSRNTIS--HIOPFLDLESLEL-----SLHLSNRRLPS 114
 Db 125 LEKHLDDNSVTSVIEDEADAFADSKQKLLFLSRNHLSSIFSGLPHTLEELRDLNRLIST 184
 Oy 115 LGEPTLRGLVNLQHLIVNNOJGG--IADEFEPLELLTLELDLSYNNLHGLPMDVSRM 172
 Db 185 IPLHAFKGLNSLRRLVLDGNLLANQRIADDTFSR--LQNLTELSLRVNSLAAPPLN--LPS 241
 Oy 173 VNLQSLDHLNLDHIABGTADLOKRLARLDLTSNRLOKLP--DPIFARSAALATAT 229
 Db 242 AHLQKLYLDNAISHI PYNTLAKRELERLRLDLSNNLTLPRGLFDDDLGDLNLAQ----- 294
 Oy 230 PFAPPLSFGNPLHNCCELLLRRLRLRDDLETCGSPGGLKGRYFWHVRREEEVCPEPP 289
 Db 295 -----LLLRNPWFCCNLMWLRD-----W----- 314
 Oy 290 LITQHTKLLVLEGOAATLKCAIGDPSPLIHVAPDRDLVGNRSRTAYVD--NGTLDIFI 348
 Db 315 -----VKARAAVVNVRGLMCOGP-----EKVRG-----MAIKDITSEMDECF 351
 Oy 349 TTSQDSGAFTCIANAAGEATAMVEVSIYQVPHLSNSTSRTP-----PKSRLS 397
 Db 352 ETGFGG-----VANAAAKTTA-----SNHASATTPQSGSLFTLAKRPGRLRP 394
 Oy 398 DITGSSKTSRGGGGGGEPPKSPPERAVLVSEVTTTSALVKWSVKSAPRVQMYQLOVN 457
 Db 395 DSNIDYPWATGDGAK-----TLAIHKVALTADSIIRITWKATLPASSFRLSGLWRIG 444
 Oy 458 CSDSEVLYRMIIPASNKA-FVYNNLVSGTYDLCVLAMWDDTATTLTATNIVGCAQFFTK 516
 Db 445 HSPAVGSITELVQGDKTEYLLTALPKSITYIICVWTM--ETSNAVVADETPVCAKAETA 502
 Oy 517 ADYPOCCQMSHSGILGGTMI-LVIGGII---VATLLVFIILMVRKYVCNHEA 564
 Db 503 DSYGPTTLNQEONAGPMASLPLAGIIGGAVALVFLVLVLAICWYV--HOA 552
 RESULT 3
 SLIT_DROME STANDARD; PRT; 1480 AA.
 AC P24014;
 DT 01-MAR-1992 (Rel. 21, Created)
 DT 01-MAR-1992 (Rel. 21, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Slit protein precursor.
 GN SLI.

OS Drosophila melanogaster (Fruit fly)
 CC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
 CC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
 CC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
 RN NCBI_TaxID=7227;
 [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91099665; PubMed=2176636;
 RA Rothberg J.M., Jacobs J.R., Goodman C.S., Artavanis-Tsakonas S.;
 RT "Slit: an extracellular protein necessary for development of midline
 RT glia and commissural axon pathways contains both EGF and LRR
 RT domains.";
 RL Genes Dev. 4:2169-2187(1990).
 CC -!- FUNCTION: NECESSARY FOR DEVELOPMENT OF MIDLINE GLIA AND
 CC COMMISSURAL AXON PATHWAYS. SLIT MAY INTERACT WITH EXTRACELLULAR
 CC MATRIX MOLECULES.
 CC -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A LONG FORM (SHOWN HERE) AND A
 CC SHORT FORM; ARE PRODUCED BY ALTERNATIVE SPLICING.
 CC -!- TISSUE SPECIFICITY: EXCRETED BY THE MIDLINE GLIA CELLS AND
 CC EVENTUALLY DISTRIBUTED ALONG THE AXONS.
 CC -!- SIMILARITY: CONTAINS 7 EGF-LIKE DOMAINS.
 CC -!- SIMILARITY: CONTAINS 24 LEUCINE-RICH REPEATS (LRR).
 CC -!- SIMILARITY: CONTAINS 1 C-TERMINAL CYSTINE KNOT-LIKE (CTCK) DOMAIN.
 CC -!- SIMILARITY: CONTAINS 1 LAMININ G-LIKE DOMAIN.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL; X53959; CAA37910.1; .
 DR PIR; A36665; A36665.
 DR HSSP; P00740; LEDM.
 DR Flybase; FBgn0003425; sli.
 DR InterPro; IPR000152; Asx_hydroxy.
 DR InterPro; IPR000359; Cys_knot.
 DR InterPro; IPR000561; EGF-like.
 DR InterPro; IPR000742; EGF_2.
 DR InterPro; IPR001881; EGF_Ca.
 DR InterPro; IPR001611; LRR.
 DR InterPro; IPR000483; LRR_Cterm.
 DR InterPro; IPR000372; LRR_Nterm.
 DR InterPro; IPR003592; LRR_out.
 DR InterPro; IPR003591; LRR_typ.
 DR InterPro; IPR001791; Laminin_G.
 DR Pfam; PF00007; Cys_knot; 1.
 DR Pfam; PF00008; EGF; 7.
 DR Pfam; PF00054; laminin_G; 1.
 DR Pfam; PF00560; LRR; 16.
 DR Pfam; PF01462; LRRNT; 4.
 DR Pfam; PF01463; LRRCT; 4.
 DR SMART; SM00041; CT; 1.
 DR SMART; SM00179; EGF_CA; 2.
 DR SMART; SM00001; EGF_like; 5.
 DR SMART; SM00370; LRR; 4.
 DR SMART; SM00082; LRRCT; 4.
 DR SMART; SM00013; LRRNT; 4.
 DR SMART; SM00369; LRR_TYP; 9.
 DR SMART; SM00282; LamG; 1.
 DR PROSITE; PS00010; ASX_HYDROXYL; 3.
 DR PROSITE; PS01185; CTCK_1; 1.
 DR PROSITE; PS01225; CTCK_2; 1.
 DR PROSITE; PS00022; EGF_1; 7.
 DR PROSITE; PS01186; EGF_2; 5.
 DR PROSITE; PS01187; EGF_CA; 2.
 DR PROSITE; PS50025; LAM_G_DOMAIN; 1.
 KW Neurogenesis; Glycoprotein; Signal; Alternative splicing;
 KW EGF-like domain; Repeat; Leucine-rich repeat.
 FT SIGNAL 1 36
 FT CHAIN 37 1480 SLIT PROTEIN.

Tue Jun 10 08:57:40 2003

99 REPEAT 122 LRR 1.
 FT REPEAT 123 LRR 2.
 FT REPEAT 146 LRR 3.
 FT REPEAT 170 LRR 4.
 FT REPEAT 171 LRR 5.
 FT REPEAT 194 LRR 6.
 FT REPEAT 218 LRR 7.
 FT REPEAT 220 LRR 8.
 FT REPEAT 246 LRR 9.
 FT REPEAT 321 LRR 10.
 FT REPEAT 345 LRR 11.
 FT REPEAT 368 LRR 12.
 FT REPEAT 392 LRR 13.
 FT REPEAT 399 LRR 14.
 FT REPEAT 416 LRR 15.
 FT REPEAT 595 LRR 16.
 FT REPEAT 617 LRR 17.
 FT REPEAT 641 LRR 18.
 FT REPEAT 643 LRR 19.
 FT REPEAT 666 LRR 20.
 FT REPEAT 701 LRR 21.
 FT REPEAT 720 LRR 22.
 FT REPEAT 743 LRR 23.
 FT REPEAT 745 LRR 24.
 FT REPEAT 764 LRR 25.
 FT REPEAT 765 LRR 26.
 FT REPEAT 788 LRR 27.
 FT REPEAT 790 LRR 28.
 FT REPEAT 812 LRR 29.
 FT REPEAT 813 LRR 30.
 FT REPEAT 836 LRR 31.
 FT REPEAT 838 LRR 32.
 FT REPEAT 861 LRR 33.
 FT REPEAT 907 LRR 34.
 FT REPEAT 944 LRR 35.
 FT DOMAIN 946 LRR 36.
 FT DOMAIN 985 LRR 37.
 FT DOMAIN 1022 LRR 38.
 FT DOMAIN 1024 LRR 39.
 FT DOMAIN 1064 LRR 40.
 FT DOMAIN 1111 LRR 41.
 FT DOMAIN 1149 LRR 42.
 FT DOMAIN 1152 LRR 43.
 FT DOMAIN 1325 LRR 44.
 FT DOMAIN 1332 LRR 45.
 FT DOMAIN 1409 LRR 46.
 FT CARBOHYD 111 LRR 47.
 FT CARBOHYD 111 LRR 48.
 FT CARBOHYD 207 LRR 49.
 FT CARBOHYD 211 LRR 50.
 FT CARBOHYD 357 LRR 51.
 FT CARBOHYD 357 LRR 52.
 FT CARBOHYD 435 LRR 53.
 FT CARBOHYD 435 LRR 54.
 FT CARBOHYD 783 LRR 55.
 FT CARBOHYD 783 LRR 56.
 FT CARBOHYD 788 LRR 57.
 FT CARBOHYD 788 LRR 58.
 FT CARBOHYD 958 LRR 59.
 FT CARBOHYD 998 LRR 60.
 FT CARBOHYD 1060 LRR 61.
 FT CARBOHYD 1060 LRR 62.
 FT CARBOHYD 1159 LRR 63.
 FT CARBOHYD 1159 LRR 64.
 FT CARBOHYD 1175 LRR 65.
 FT CARBOHYD 1175 LRR 66.
 FT CARBOHYD 1243 LRR 67.
 FT CARBOHYD 1243 LRR 68.
 FT CARBOHYD 1292 LRR 69.
 FT CARBOHYD 1292 LRR 70.
 FT DISULFID 911 LRR 71.
 FT DISULFID 911 LRR 72.
 FT DISULFID 916 LRR 73.
 FT DISULFID 932 LRR 74.
 FT DISULFID 934 LRR 75.
 FT DISULFID 943 LRR 76.
 FT DISULFID 950 LRR 77.
 FT DISULFID 950 LRR 78.
 FT DISULFID 955 LRR 79.
 FT DISULFID 955 LRR 80.
 FT DISULFID 973 LRR 81.
 FT DISULFID 973 LRR 82.
 FT DISULFID 989 LRR 83.
 FT DISULFID 989 LRR 84.
 FT DISULFID 995 LRR 85.
 FT DISULFID 995 LRR 86.
 FT DISULFID 1012 LRR 87.
 FT DISULFID 1012 LRR 88.
 FT DISULFID 1028 LRR 89.
 FT DISULFID 1028 LRR 90.
 FT DISULFID 1035 LRR 91.
 FT DISULFID 1035 LRR 92.
 FT DISULFID 1052 LRR 93.
 FT DISULFID 1052 LRR 94.
 FT DISULFID 1068 LRR 95.
 FT DISULFID 1068 LRR 96.
 FT DISULFID 1073 LRR 97.
 FT DISULFID 1073 LRR 98.
 FT DISULFID 1090 LRR 99.
 FT DISULFID 1090 LRR 100.
 FT DISULFID 1115 LRR 101.
 FT DISULFID 1115 LRR 102.
 FT DISULFID 1125 LRR 103.
 FT DISULFID 1125 LRR 104.
 FT DISULFID 1137 LRR 105.
 FT DISULFID 1137 LRR 106.
 FT DISULFID 1148 LRR 107.
 FT DISULFID 1148 LRR 108.
 FT DISULFID 1357 LRR 109.
 FT DISULFID 1357 LRR 110.
 FT DISULFID 1362 LRR 111.
 FT DISULFID 1362 LRR 112.
 FT DISULFID 1380 LRR 113.
 FT DISULFID 1380 LRR 114.
 FT DISULFID 1391 LRR 115.
 FT DISULFID 1391 LRR 116.
 FT DISULFID 1409 LRR 117.
 FT DISULFID 1409 LRR 118.
 FT DISULFID 1423 LRR 119.
 FT DISULFID 1423 LRR 120.
 FT DISULFID 1434 LRR 121.
 FT DISULFID 1434 LRR 122.
 FT DISULFID 1438 LRR 123.
 FT DISULFID 1438 LRR 124.
 FT DISULFID 1442 LRR 125.
 FT DISULFID 1442 LRR 126.
 FT VARSPLIC 1394 LRR 127.
 FT VARSPLIC 1394 LRR 128.

SQ SEQUENCE 1480 AA; 165752 MW; P9DS925FC170B1C3 CRC64;
 Query Match 6.4%; Score 263; DB 1; Length 1480;
 Best Local Similarity 28.0%; Pred. No. 3.7e-09;
 Matches 89; Conservative 39; Mismatches 102; Indels 88; Gaps 11;
 QY 6 GGLLAFCMAF-----AVVDA-CPKVCYCVONLSESLGTLCPKSGLLFVFPDPIDRRITVELR 58
 DB 51 GGLGVGIHPIGGGVGVITFARCPVCSCTGLNVD---CSHRLTSVPRKISADYVERLE 106
 QY 59 LGNFIHISRQDFANMTGLVTLTSGRTISHTISHIPQFSFDFLDELRSRHLSDNRLPSLGED 118
 DB 107 LOGNLTVIYETDFQRLTKRLMLQLTDNQIHTIERNSTQDLVSLERLDISNNVTITVGR 166
 QY 119 TLRGLVNLQHLIYNNQLGGADEAFEDLTDLDELNLLHGLPWSVRRMVLHQL 178
 DB 167 VFKGAQSLRSLQDNNQITCDDEHAFKG-LVELEILTNNNTLSTLP----- 212
 QY 179 SLDHNLHDIAEGTFADLQKLARDLTLNRLQKLPDPPIFARSOASALTAPPAPLUSFS 238
 DB 213 ---HNI-----FGGLGRALRLSDN-----PFA----- 233
 QY 239 FGNPLHCNCELLWRLRLRD-----DDLETCSGSPGGLKGRYFVHVRREEEFCBPPLITOH 294
 DB 234 -----CDGHLNLSRFLRSATRLAPYTRCQSPSLKQGVADLHDQEFKCSG--LIEH 284
 QY 295 THKLVLVEGOAATLKCKA 312
 DB 285 -----APMECGA 291
 RESULT 4
 ID_ ALS RAT STANDARD; PRT; 603 AA.
 AC P35859;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Insulin-like growth factor binding protein complex acid labile chain precursor (ALS)
 GN IGFALS OR ALS.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
 ON NCBI TaxID=10116;
 RX [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RX MEDLINE=93038676; PubMed=1384485;
 RA Dai J., Baxter R.C.;
 RT "Molecular cloning of the acid-labile subunit of the rat insulin-like growth factor binding protein complex."
 RL Biochem. Biophys. Res. Commun. 188:304-309(1992).
 [2]
 RP SEQUENCE OF 24-44, AND CHARACTERIZATION.
 RC STRAIN=Wistar; TISSUE=Sera;
 RX MEDLINE=94130835; PubMed=7507839;
 RA Baxter R.C., Dai J.;
 RT "Purification and characterization of the acid-labile subunit of rat serum insulin-like growth factor binding protein complex."
 RL Endocrinology 134:848-852(1994).
 CC -!- FUNCTION: MAY HAVE AN IMPORTANT ROLE IN REGULATING THE ACCESS OF
 CC -!- SUBUNIT: FORMS A TERNARY COMPLEX OF ABOUT 140 TO 150 kDa WITH
 CC IGF-1 OR IGF-II AND IGFBP-3.
 CC -!- SUBCELLULAR LOCATION: Extracellular.
 CC -!- TISSUE SPECIFICITY: BRAIN, KIDNEY, LUNG, HEART, SPLEEN, MUSCLE
 CC AND LIVER.
 CC -!- SIMILARITY: CONTAINS 21 LEUCINE-RICH REPEATS (LRR).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch)
 CC -----

DR EMBL; S46785; AAR23770.2; .
 DR PIR; JCL1282; JCL1282. LRR.
 DR InterPro; IPR001611; LRR.
 DR InterPro; IPR000483; LRR Cterm.
 DR InterPro; IPR000372; LRR Nterm.
 DR InterPro; IPR003592; LRR out.
 DR InterPro; IPR003591; LRR typ.
 DR Pfam; PF00560; LRR; 18.
 DR Pfam; PF01462; LRRNT; 1.
 DR Pfam; PF01463; LRRCT; 1.
 DR PRINTS; PRO0019; LEURICHRPT.
 DR SMART; SM00370; LRR; 5.
 DR SMART; SM00082; LRRCT; 1.
 DR SMART; SM00013; LRRNT; 1.
 DR SMART; SM00369; LRR_TYP; 9.
 KW Glycoprotein; Leucine-rich repeat; Repeat; Signal.
 FT SIGNAL 1 23
 FT CHAIN 24 603
 FT REPEAT 52 73 INSULIN-LIKE GROWTH FACTOR BINDING
 FT REPEAT 74 96 LRR 1.
 FT REPEAT 98 120 LRR 2.
 FT REPEAT 121 144 LRR 3.
 FT REPEAT 146 168 LRR 4.
 FT REPEAT 169 192 LRR 5.
 FT REPEAT 194 216 LRR 6.
 FT REPEAT 217 240 LRR 7.
 FT REPEAT 242 264 LRR 8.
 FT REPEAT 266 288 LRR 9.
 FT REPEAT 289 312 LRR 10.
 FT REPEAT 313 336 LRR 11.
 FT REPEAT 337 360 LRR 12.
 FT REPEAT 361 384 LRR 13.
 FT REPEAT 386 408 LRR 14.
 FT REPEAT 409 432 LRR 15.
 FT REPEAT 433 456 LRR 16.
 FT REPEAT 458 478 LRR 17.
 FT REPEAT 479 504 LRR 18.
 FT REPEAT 506 529 LRR 19.
 FT REPEAT 543 566 LRR 20.
 FT CARBOHYD 64 64 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 85 85 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 96 96 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 368 368 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 515 515 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 578 578 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 586 586 N-LINKED (GLCNAC. .) (POTENTIAL).
 SQ SEQUENCE 603 AA; 66811 MW; DCD7637D94A5037C CRC64;

Query Match 6.1%; Score 248; DB 1; Length 603;
 Best Local Similarity 25.2%; Pred. No. 1e-08;
 Matches 104; Conservative 64; Mismatches 161; Indels 84; Gaps 15;
 OY 8 LLAFGMAF-----AVVDA---CPKVCVC--QNLSESLGLCPKGLLFPVDDI 50
 DB 13 LLAFWALGPHCLOGTDFGASADAEQPQCVACTCSHDDYDDELVSFCSSKNLTHLPDDI 72
 OY 51 DRRTVELRGGNFIIHQFANMTGLVDLTLRSNTISHIOPFFSLDLESLSLHDSN 110
 DB 73 PVSTRALWLDGNLSSISPAFQNLSSLDLFLNQLQSWLRSLEPQALLGLQNLVYLHLSRN 132
 OY 111 RLPSLGEDTLRGLVNLQHLVNNQGGIADBAFEDFLTLLEDLDSYNNLHGLPMDSVR 170
 DB 133 RLRLNLAUGLFTHTPSLASLSLSSNLLGRLEGLFQGL--LSHLWDLNLGWNLSLWLPDPTVQ 191
 OY 171 RMVNLHQLSLDHLNLDHAEFTFADLQKLRDLTNSLQKLPDPPDFARQASALATFP 230
 DB 192 GLGNLHELVLVAGNKLTLYLQPALFCGLBELRELDLSRNALRSV-----KANVFVHLP 242

OY 231 FAPPLSFSFGNPLHCNCELMLRRLERDDLETCGSPGLKG--RYPW-----HVR----- 280
 DB 243 -----RLQKLYLDRLNLTAVAPGAFGLMKALRWLDSLHNRVAGL 281
 OY 281 -EEBFVCEPPLITHTHKLLEGOAATLKCAIKGDPSPPLJHWVAPDRDLVGNSSRTAVY 339
 DB 282 MEDTF---PGLLGLHV--LRLAHNAIASLRPTFKD-----LHFL--EELQGHNRIRQLG 330
 OY 340 DN-----GTLDFI-----ITTSQDSGAFTCIAANAAAGEATAMVFEVSIQ 378
 DB 331 ERTEFGLQLEVLTLNDNQITEVRVGFAPSGLFNVAVMNLSGNCLSLRFERVFQ 383

RESULT 5

CHAD_HUMAN
 ID CHAD_HUMAN STANDARD; PRT; 359 AA.
 AC O15335; Q96RJ5;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DE 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Chondroadherin precursor (Cartilage leucine-rich protein).
 GN CHAD.
 OS Homo sapiens (Human)
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A. PubMed=9344663;
 RX MEDLINE=98008928; PubMed=9344663;
 RA Grover J., Chen X.-N., Korenberg J.R., Roughley P.J.;
 RT "The structure and chromosome location of the human chondroadherin
 gene (CHAD).";
 RL Genomics 45:379-385(1997).
 RN [2]
 RP SEQUENCE FROM N.A. PubMed=11415564;
 RX MEDLINE=21413956; PubMed=11415564;
 RA Maansson B., Wenglen C., Moergelin M., Saxne T., Heinegaard D.;
 RT "Association of chondroadherin with collagen type II.";
 RL J. Biol. Chem. 276:32883-32888(2001).
 CC !- FUNCTION: Promotes attachment of chondrocytes, fibroblasts, and
 CC osteoblasts. This binding is mediated (at least for chondrocytes
 CC and fibroblasts) by the integrin alpha(2)beta(1). May play an
 CC important role in the regulation of chondrocyte growth and
 CC proliferation (By similarity).
 CC !- SUBUNIT: Mostly monomeric (By similarity). Interacts with collagen
 CC type II.
 CC !- SUBCELLULAR LOCATION: Extracellular matrix (By similarity).
 CC !- TISSUE SPECIFICITY: Present in chondrocytes at all ages.
 CC !- SIMILARITY: BELONGS TO THE SMALL LEUCINE-RICH PROTEOGLYCAN
 CC (SLRPS) FAMILY. CLASS IV SUBFAMILY.
 CC !- SIMILARITY: CONTAINS 11 LEUCINE-RICH REPEATS (LRR).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch)
 CC -----
 DR EMBL; U96769; AAC13410.1; .
 DR EMBL; U96767; AAC13410.1; JOINED.
 DR EMBL; U96768; AAC13410.1; JOINED.
 DR EMBL; AF371328; AAK51556.1; .
 DR Genew; HGNC:1909; CHAD.
 DR MIM; 602178; .
 DR InterPro; IPR001611; LRR.
 DR InterPro; IPR000483; LRR Cterm.
 DR InterPro; IPR000372; LRR Nterm.
 DR InterPro; IPR003592; LRR out.
 DR InterPro; IPR003591; LRR typ.
 DR Pfam; PF00560; LRR; 10.

QY 322 WVPDDRVLVGNSSRTAVYDN--GTLDFITTS-----ODSAGFTC----- 359
 Db 349 RVRPTSR--OPSKDLGYNIGPSYAYQTKSPVLECPCTACSCNQLQISDLGLNVCQERK 405
 QY 360 -----IAAN--AAGATAMVEYSIVQLPHLSNSTRTAPPKSR-LSD 398
 Db 406 IESIAELQPKPNPKMYLNTENYIAVVRRTDFLEATGLDHLHGN--NRISMIOQRAFGD 463
 QY 399 ITGSSKTSRGGGGGGBEPKSPPERAVLSEVITTSALVKWSVSKSAPR-----VKM 451
 Db 464 LTNLRLRYLNGN-----RIEELPELFGLOSLOY 493
 QY 452 YOLQVN--CSDDEVLVYRMPASNKAFVNNLV-----SG--TGYDLCVLAMWDDTATTLTA 504
 Db 494 LFLQVNLRIEIQSGTFDPVNLQLLFLNLLQAMPSGVFSLGTLRLNLRNHNHTSLPV 553
 QY 505 TNIVGCAQFFKADYQPCQSMHSQILGGMILVIGGIIIVATLLVIVILVYKVCNHEA 564
 Db 554 SGVLDQLKSLIQID-----LHDNPMWCTCDIVGMKLVWEQLKGVLVDEV---IC--KA 602
 QY 565 PSMVAANVSNVYQTNQAQPPSSAPAGAPQGGPKVYVVRNELLDFEASLARASDSSSS 624
 Db 603 PKKFAE--TDMRSIKSELCPDYSVVVSTPSSIQVPARTSAV---TPAVRLNSTGAP 657
 QY 625 SSLGSGEAGLGRAPWRIPPSAPRPKSLDRMLMGAPAS-----LDLKSQRKEELDSRT- 678
 Db 658 ASLGAGGGAS-----SVPLSVLILSLLLVFMSVFAAGLVLMKRRKKNQSDHTSTN 711
 QY 679 -----PAGCAGTARSRHSDREPLL---GPPAARARSLPLPL 714
 Db 712 NSDVSSFNQVSVYGGGGTGGHPAHVHRGPPALPKVKTAPAGHYVEYIPIHPL 764

RESULT 7
 ALS_HUMAN STANDARD; PRT; 605 AA.
 AC P35858;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DE Insulin-like growth factor binding protein complex acid labile chain precursor (ALS).
 GN IGFALS OR ALS.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RC TISSUE=Liver;
 RA Leong S.R., Baxter R.C., Camerato T., Dai J., Wood W.I.;
 RT "Structure and functional expression of the acid-labile subunit of the insulin-like growth factor-binding protein complex.";
 RL Moi. Endocrinol. 6:870-876(1992).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Frankland J.;
 RL Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE OF 28-35.
 RX MEDLINE=89308584; PubMed=2473065;
 RA Baxter R.C., Martin J.L., Beniac V.A.;
 RT "High molecular weight insulin-like growth factor binding protein complex. Purification and properties of the acid-labile subunit from human serum.";
 RL J. Biol. Chem. 264:11843-11848(1989).
 CC -!- FUNCTION: INVOLVED IN PROTEIN-PROTEIN INTERACTIONS THAT RESULT IN PROTEIN COMPLEXES, RECEPTOR-LIGAND BINDING OR CELL ADHESION.
 CC -!- SUBUNIT: FORMS A TERNARY COMPLEX OF ABOUT 140 TO 150 kDa WITH IGF-I OR IGF-II AND IGFEBP-3.
 CC -!- SUBCELLULAR LOCATION: Extracellular.

CC -!- TISSUE SPECIFICITY: PLASMA.
 CC -!- SIMILARITY: CONTAINS 20 LEUCINE-RICH REPEATS (LRR).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M86826; AAA36047.1; -;
 DR EMBL; AL031724; CAC36078.1; -;
 DR PIR; A41915; A41915.
 DR Genew; HGNC:5468; IGFALS.
 DR MIM; 601489; -;
 DR InterPro; IPR001611; LRR.
 DR InterPro; IPR000483; LRR_Cterm.
 DR InterPro; IPR000372; LRR_Nterm.
 DR InterPro; IPR003592; LRR_out.
 DR InterPro; IPR003591; LRR_typ.
 DR Pfam; PF00560; LRR; 19.
 DR Pfam; PF01462; LRRNT; 1.
 DR Pfam; PF01463; LRRCT; 1.
 DR PRINTS; PR00019; LEUKICHRPT.
 DR SMART; SM00370; LRR; 2.
 DR SMART; SM00082; LRRCT; 1.
 DR SMART; SM00013; LRRNT; 1.
 DR SMART; SM00369; LRR_TYP; 11.
 KW Glycoprotein; Leucine-rich repeat; Repeat; Signal.
 FT SIGNAL 1 27
 FT CHAIN 28 605
 FT REPEAT 53 73
 FT REPEAT 74 96
 FT REPEAT 98 120
 FT REPEAT 121 144
 FT REPEAT 145 168
 FT REPEAT 170 192
 FT REPEAT 193 216
 FT REPEAT 217 240
 FT REPEAT 242 264
 FT REPEAT 265 288
 FT REPEAT 289 312
 FT REPEAT 313 336
 FT REPEAT 337 360
 FT REPEAT 361 384
 FT REPEAT 386 408
 FT REPEAT 409 432
 FT REPEAT 433 456
 FT REPEAT 458 480
 FT REPEAT 482 504
 FT REPEAT 505 530
 FT CARBOHYD 64 64
 FT CARBOHYD 85 85
 FT CARBOHYD 96 96
 FT CARBOHYD 368 368
 FT CARBOHYD 515 515
 FT CARBOHYD 580 580
 SQ SEQUENCE 605 AA; 66034 MW; F6562A23CBE918F6 CRC64;

Query Match 5.9%; Score 241.5; DB 1; Length 605;
 Best Local Similarity 29.7%; Pred. No. 2.7e-08;
 Matches 76; Conservative 41; Mismatches 106; Indels 33; Gaps 4;
 QY 20 ACPKYVCV--QNLSESLGTLCPKSGLLFVPPDIDRTVELRGGNFIIHSRODFANMTG 77
 Db 40 ACPAAVCVSYDDADELSEVFCSSRNLRPLPDGVPDGTQALWLDGNNLSSVPPAATONLSS 99
 QY 78 LVDLTLRNTTISHQPFPSFLDESRLHDSNRLPSLIGEDTLRGLVNLQHLVNNQIG 137
 Db 100 LGLNQLQGGQLGSEPPQALLGLELHLEHNRNQLRSLALGTFHTPALASLGLSNNRLS 159

QY 138 GIADAEFDFLLTLEDDLSYNNHLCPWDSVRRMVLHQLSDHNLDDHIAEGTFADLQ 197
 Db 160 RLEDGLFEG-IGSLWDLNLGWNLSLAVLDPDAFPGLSRELVLGNRLAYLQPALFSGLA 218
 QY 198 KIARLDLTSNRIOKLPDPIFARSOASALATFPFAPPPLSFSGNPLHCNCELLWLRLLE 257
 Db 219 ELRELDLGRNALR-----AIKANVFV-----QLPRLOKLY 248
 QY 258 RDDDLETGCPGLK 273
 Db 249 LDRNLIAAVAPGAFILG 264

RESULT 8

FLR3_HUMAN STANDARD; PRT; 649 AA.
 AC Q9NZU0; Q96KB1; Q9P259; Q96K42; Q96K39;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Leucine-rich repeat transmembrane protein FLRT3 precursor
 (Fibronectin-like domain-containing leucine-rich transmembrane protein
 3).
 GN FLRT3 OR KIAA1469.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.
 RX MEDLINE=20112755; PubMed=10644439;
 RA Lacy S.E., Bonnemant C.G., Ruzney E.A., Kunkel L.M.;
 RT "Identification of FLRT1, FLRT2, and FLRT3: a novel family of
 transmembrane leucine-rich repeat proteins.";
 RL Genomics 62:417-426(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=20277482; PubMed=10819331;
 RA Nagase T., Kikuno R., Ishikawa K.-I., Hiroswawa M., Ohara O.;
 RT "prediction of the coding sequences of unidentified human genes. XVII.
 The complete sequences of 100 new cDNA clones from brain which code
 for large proteins in vitro.";
 RL DNA Res. 7:143-150(2000).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Embryo, and Teratocarcinoma;
 RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
 RA Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,
 RA Wagatsuma M., Hosoi T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
 RA Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,
 RA Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,
 RA Yamamoto J., Wakamatsu A., Nakamura Y., Nagahari K., Masuho Y.,
 RA Ninomiya K., Iwayanagi T.;
 RT "NEDO human cDNA sequencing project.";
 RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21638749; PubMed=11780052;
 RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,
 RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Baggeley C.L.,
 RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,
 RA Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,
 RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,
 RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,
 RA Clegg S., Copley V.E., Collier R.E., Connor R.E., Corby N.R.,
 RA Coulson A., Coville G.J., Deadman R., Dhami P.D., Dunn M.,
 RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,
 RA Grafham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,
 RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,
 RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,
 RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,
 RA Lechvasiaino M.H., Leversha M.A., Lloyd C., Lloyd D.M., Lovell J.D.,

RA Marsh V.L., Martin S.L., McConnachie L.J., Mcclay K., McMurray A.A.,
 RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,
 RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,
 RA Phillimore B.J.C.T., Prachalingam S.R., Plumb R.W., Ramsay H.,
 RA Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Showkeen R., Sims S.,
 RA Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,
 RA Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,
 RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M., Williams S.A.,
 RA Whitehead S.L., Whittaker P., Willey D.L., Williams L., Williams D.R.,
 RA Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,
 RA Rogers J.;
 RT "The DNA sequence and comparative analysis of human chromosome 20.";
 RL Nature 414:865-871(2001).
 RN [5]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Kidney;
 RA Strausberg R.;
 RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: May have a function in cell adhesion and/or receptor
 signaling.
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein (Probable).
 CC -!- TISSUE SPECIFICITY: Expressed in kidney, brain, pancreas, skeletal
 muscle, lung, liver, placenta, and heart.
 CC -!- SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE III-LIKE DOMAIN.
 CC -!- SIMILARITY: CONTAINS 10 LEUCINE-RICH REPEATS (LRR).
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (see http://www.isb-sib.ch/announce/
 or send an email to license@sib-sib.ch).
 DR EMBL; AF169677; AAF28461.1; -
 DR EMBL; AB040902; BAA95993.1; ALT_INIT.
 DR EMBL; AK027297; BAB55023.1; -
 DR EMBL; AK027670; BAB55282.1; ALT_INIT.
 DR EMBL; AK027694; BAB55303.1; ALT_INIT.
 DR EMBL; AL132826; CAB86687.1; -
 DR EMBL; BC020870; AAH20870.1; -
 DR Genew; HGNC:3762; FLRT3.
 DR MIM; 604808; -
 DR InterPro; IPR003961; FN_III.
 DR InterPro; IPR001611; LRR.
 DR InterPro; IPR000483; LRR_Cterm.
 DR InterPro; IPR000372; LRR_Nterm.
 DR InterPro; IPR003592; LRR_out.
 DR InterPro; IPR003591; LRR_typ.
 DR Pfam; PF00041; In3; 1.
 DR Pfam; PF00560; LRR; 10.
 DR Pfam; PF01462; LRRNT; 1.
 DR Pfam; PF01463; LRRCT; 1.
 DR SMART; SM00060; FN3; 1.
 DR SMART; SM00370; LRR; 1.
 DR SMART; SM00082; LRRCT; 1.
 DR SMART; SM00013; LRRNT; 1.
 DR SMART; SM00369; LRR_TYP; 1.
 DR Transmembrane; Leucine-rich repeat; Repeat; Glycoprotein; Signal.
 DR SIGNAL
 DR CHAIN 1 28
 FT SIGNAL 29 649
 FT LEUCINE-RICH REPEAT TRANSMEMBRANE PROTEIN
 FT POTENTIAL.
 FT FLRT3
 FT EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 529 549
 FT DOMAIN 550 649
 FT CYTOPLASMIC (POTENTIAL).
 FT REPEAT 57 82
 FT LRR 2.
 FT REPEAT 83 105
 FT LRR 3.
 FT REPEAT 107 126
 FT LRR 4.
 FT REPEAT 127 152
 FT LRR 5.
 FT REPEAT 154 179
 FT LRR 6.
 FT REPEAT 181 197
 FT LRR 7.
 FT REPEAT 198 223
 FT LRR 8.
 FT REPEAT 224 246

FT REPEAT	247	269	LRR 9.	
FT REPEAT	270	485	LRR 10.	
FT DOMAIN	405	293	FIBRONECTIN TYPE-III.	
FT CARBOHYD	226	282	N-LINKED (GLCNAC. . .)	(POTENTIAL)
FT CARBOHYD	282	282	N-LINKED (GLCNAC. . .)	(POTENTIAL)
FT CARBOHYD	296	296	N-LINKED (GLCNAC. . .)	(POTENTIAL)
FT CONFLICT	198	198	L -> P (IN REF. 3; BABS5282).	
FT CONFLICT	307	307	W -> R (IN REF. 3; BABS5282).	
FT CONFLICT	400	400	H -> Q (IN REF. 3; BABS5023).	
FT CONFLICT	519	519	K -> Q (IN REF. 3; BABS5023).	
FT CONFLICT	638	638	D -> G (IN REF. 3; BABS5282).	
SQ SEQUENCE	649 AA;	73003 MW;	9EFF666C46181F08 CRC64;	
Query Match		5.9%;	Score 240.5; DB 1; Length 649;	
Best Local Similarity		21.8%;	Pred. No. 3.4e-08;	
Matches 141;		Conservative 90;	Mismatches 220; Indels 195; Gaps 28;	
Qy	3	TLLGGLAFGMFAVVDACPKYCVCO	-----NLSESLGTL-----	37
Db	13	TKIGLFLQVAPUSVMAKSCPSVCRDAGFIYCNDRFLTSIPTGIPEDATTLYLQNNQINN	72	
Qy	38	--CPS--KGLLFV-----	PPDIDRRTVELRLGGNFIIHISRODFANMTGLVD	80
Db	73	AGIPSDLKMLKVERIYLYHNSLDFEFTNPKYVKELHLQENNIRITYVDSLSKIPYLEE	132	
Qy	81	L7LSRNTS--HIQPFSLDLESR-----	SLHLSNRLPFSLGE	117
Db	133	LHLDNSVSAVSIERGAFRDSNYLRLLFLSRNHLSTIPWGLPRTIEELRLDDNRISTISS	192	
Qy	118	DTLRGLVNLHLIVNNOLG--GIADAEFEDLLEDLDSYNLHLGLPWSVRMVNL	175	
Db	193	PSLQGLSLKRLVLDGNLNNHGLGDKVFFN-LVNLTELSLVRNSLTAAPVNLPG--TNL	249	
Qy	176	HOLSLDHLLDHAEGTTFADLOKLARLDLTSNRLOKLPDPDFARSOASALATAPFAPPL	235	
Db	250	RKLYLQDNHINRVPPNAFSYLRQLYELDMNSNLSNL-PQGIFF--DDLDNIT-----	298	
Qy	236	SFSFGNPLHMCCELLLWRLRLRERDDLET-----	CGSPGLKGRYFMVHREBEF	284
Db	299	QLILRNPNWYCCCKMKWR-----	DWLQSLPVKVVNRGLMCOAPEKVRGMAIKDLNAELF	353
Qy	285	VCEP-----PLIOTHHLKLVLSGO--AATLKCAIGDPSPLHWAPDDRLVGNSSR	335	
Db	354	CKDGSIVSTIGITTAIPNTVYPAQGWPAVTKQPDIKNPK-----	LTKDHOITGSPSR	408
Qy	336	TAVYDNGTLDIITTSQ-----	DSGAFCTAAN-----	363
Db	409	KTI-----TIVKSVTSDITHISWKLALPMTALRSLWKLGHSPAGSITETIVTGERSEY	464	
Qy	364	--AGEATAMVEVSIQLP-----HLSNST-----	SRTAPPKSRLEDITGSSKTSRGGGGS	412
Db	465	LVTALEPDPSPYKVCMPMETSNLVYLFDETPVCIEETAPLRWYNTTTLNREQEK-----	519	
Qy	413	GGGEPKPS--PERAVLSEVTTTS-----	ALYKMSV-----SKSAPRVKMYQ	453
Db	520	---EPYKPNPLPAALIGGVALVTTIALLALVCWVYHRNGLSFRNCAYSKGRRRKDDYA	576	
Qy	454	LOYNGSDDEVL-----IYRMIPASN-----	KAFVNNLVSGTGYDL	489
Db	577	EAGTKKDSILIRETSFQMLPISNEPISKEEFVHTIPFPNGMNL	622	
RESULT 9				
LGR4 RAT				
ID	LGR4 RAT			
AC	Q9Z2H4;	STANDARD;	PRT;	951 AA.
DT	15-JUN-2002 (Rel. 41, Created)			
DI	15-JUN-2002 (Rel. 41, Last sequence update)			
DE	15-JUN-2002 (Rel. 41, Last annotation update)			
GN	GPR48 OR LGR4			
OS	Rattus norvegicus (Rat).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

NCBI_TaxID=10116;

[1]

SEQUENCE FROM N.A.

TISSUE=Ovary;

MEDLINE=99065210; PubMed=9849958;

Hsu S.Y., Liang S.-G., Hsueh A.J.W.;

"Characterization of two LGR genes homologous to gonadotropin and thyrotropin receptors with extracellular leucine-rich repeats and a G protein-coupled, seven-transmembrane region.";

Mol. Endocrinol. 12:1830-1845(1998).

CC !- FUNCTION: Orphan receptor.

CC !- SUBCELLULAR LOCATION: Integral membrane protein.

CC !- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.

CC !- SIMILARITY: CONTAINS 15 LEUCINE-RICH REPEATS (LRR).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; AF061443; AAC77910.1; -

DR InterPro; IPR000276; GPCR_Rhodpsn.

DR InterPro; IPR001611; LRR.

DR InterPro; IPR00372; LRR Nterm.

DR InterPro; IPR003592; LRR out.

DR InterPro; IPR003591; LRR_type.

DR Pfam; PF00001; 7tm_1; 1.

DR Pfam; PF00560; LRR; 15.

DR PRINTS; PR00237; GPCRHHODOPSN.

DR PRINTS; PR00019; LEURICHRPT.

DR SMART; SM00370; LRR; 4.

DR SMART; SM00013; LRRNT; 1.

DR SMART; SM00369; LRR_TYP; 5.

DR PROSITE; PS00237; G-PROTEIN RECP_F1_1; FALSE_NEG.

DR PROSITE; PS0262; G-PROTEIN RECP_F1_2; 1.

KW G-protein coupled receptor; Transmembrane; Signal; Glycoprotein;

KW Repeat; Leucine-rich repeat.

FT SIGNAL 1 24

FT CHAIN 25 951

FT FT 25 544

FT DOMAIN 25 544

FT TRANSMEM 545 565

FT DOMAIN 566 575

FT TRANSMEM 576 596

FT DOMAIN 597 619

FT TRANSMEM 620 640

FT DOMAIN 641 661

FT TRANSMEM 662 682

FT DOMAIN 683 703

FT TRANSMEM 704 724

FT DOMAIN 725 756

FT TRANSMEM 757 777

FT DOMAIN 778 783

FT TRANSMEM 784 804

FT DOMAIN 805 951

FT REPEAT 55 79

FT REPEAT 81 103

FT REPEAT 104 127

FT REPEAT 128 151

FT REPEAT 152 175

FT REPEAT 176 199

FT REPEAT 200 223

FT REPEAT 224 247

FT REPEAT 248 270

FT REPEAT 271 294

FT REPEAT 295 318

FT REPEAT 319 342

FT REPEAT 343 366

FT REPEAT 367 387

FT REPEAT 388 411 LRR 14.
 FT REPEAT 413 435 LRR 15.
 FT DISULFID 618 693 BY SIMILARITY.
 FT CARBOHYD 68 69 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 188 188 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 199 199 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 294 294 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 314 314 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 505 505 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 951 AA; 104138 MW; EDD56AC072123461 CRC64;

Query Match 5.9%; Score 240.5; DB 1; Length 951;
 Best Local Similarity 22.3%; Pred. No. 5.7e-08;
 Matches 152; Conservative 84; Mismatches 230; Indels 217; Gaps 29;

QY 7 GLLAF-----GMAFAVVDACPKYCVQNLSESLGTLCPKSGLLFVPP----- 48
 DB 6 GLLCFLLGLLGSAGSPGAAPLCAAPCSGDG---DRRVDCSGKGLTAVPEGLSAFTQAL 62
 QY 49 DDIRTV-----ELRGGNFIIHSRQDFANWTGLVDLTLSENTHSHIQP 93
 DB 63 DISMNNITQLPEDAFKSPFLEELQAGNDLSLHPKLSGLKELKVLTLQNNQRLTVP 122
 QY 94 FSPDLSELRSLHLDNSRPSLGEDTLRGLVNLQHLIVNNNOGLGGIADAFEDFLTLLED 153
 DB 123 EAIHGLSALQSRLDANHITSVPEDSFEGLVQLRHLWLDNDSITEVYRPLSN-LPTLQA 181
 QY 154 LDLSYNNLHGLPWSVRRMNLHQLSDHNLHDHAEFTADLQKRLARLDLTSNRLOKLP 213
 DB 182 LTLALNNSIPDFATNLSLVLHLHNNKIKLSLQHCDFGLDNLDTLDDNLYNLDPEP 241
 QY 214 PDFIFARSONASALTATFPAPPLSF-----SFGGNPLHNCCELLWLRRLERDDDL 262
 DB 242 -----QAIKALPSLKELFHENSISVIPDGFAGGNPL-----LRTIHLVD-- 281
 QY 263 ETCGSPGLKGRYFMH-----VREBEFVCEPLLIQTHTH-KLLVLEGGQAALTKCA 312
 DB 282 ---NPLSFGNSAFHNSLDLHCLVIRGASLVQWFFNLGTCTVHLESLLTG---TKISS 333
 QY 313 IGDPSLIHWVADDRILVGNSSRTAVYDNGTLDIFITTSQDSGAFICIAANRAGEATMV 372
 DB 334 I-----PDD-LCQKQMLR-----TLDSYNNIRDLPSF-----NGCRALE 368
 QY 373 EVSIV-OLPHLSNSTRAPPKSRSLDITGSKTSRGGGGGGGPPKSPPERAVLVE 430
 DB 369 EISLQRNQLIKENTFO-GLTSRLDLRSN-----LIRE 403
 QY 431 V-----TTTSALYKWSVKSAPR-----VKMYQLQVNCSDDEVLIYRMIPAKNAF 476
 DB 404 IHGAPAKLGTIINLDVSNFELTSFPTEGLNGLNQLKLVGNFKLKDAL-----AARDFA 457
 QY 477 VVNLVSGTGYDLCLVLAAMDATTTLTATNIVGC--AQFFTKADYPOCOSM----- 525
 DB 458 NLRSLSPVYAYQCC--APW-----GCDSYANLNTEDNSPQESHVTKKGA 502
 QY 526 -----HSOI-----LGGTMI-LVIGGIIVATLLVFIILMV 555
 DB 503 AANVTSTAEHEHSQIILHCTPSGAFKPCPEYLLGSWMIRLTVWFIFVALLFNLLVILT 562
 QY 556 RYKVCNHEAPSKM---AAAVSNV 575
 DB 563 VFASCSSLPASKLFIGLISVSNL 585

RESULT 10
 FLR2 HUMAN STANDARD; PRT; 660 AA.
 AC O43155;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Leucine-rich repeat transmembrane protein FLR2 precursor
 DE (Fibronectin-like domain-containing leucine-rich transmembrane protein

DE 2).
 GN Homo sapiens (Human).
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN (1)
 RP SEQUENCE FROM N.A.; TISSUE SPECIFICITY, AND GLYCOSYLATION.
 RX MEDLINE=20112755; PubMed=10644439;
 RA Lacy S.E., Bonemann C.G., Buzney E.A., Kunkel L.M.;
 RT "Identification of FLR1, FLR2, and FLR3: a novel family of
 RL transmembrane leucine-rich repeat proteins.";
 RL Genomics 62:417-426(1999).
 RN (2)
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=98116655; PubMed=9455477;
 RA Ishikawa K.-I., Nagase T., Nakajima D., Seki N., Ohira M.,
 RT Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;
 RT "Prediction of the coding sequences of unidentified human genes. VIII.
 RT 78 new cDNA clones from brain which code for large proteins in
 RT vitro.";
 RL DNA Res. 4:307-313(1997).
 CC -!- FUNCTION: May have a function in cell adhesion and/or receptor
 CC signaling.
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein (Probable).
 CC -!- TISSUE SPECIFICITY: Expressed in pancreas, skeletal muscle, brain,
 CC and heart.
 CC -!- PTM: N-glycosylated.
 CC -!- SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE III-LIKE DOMAIN.
 CC -!- SIMILARITY: CONTAINS 10 LEUCINE-RICH REPEATS (LRR).
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; AF169676; AAF28460.1; -;
 DR EMBL; AB007865; BAA23701.1; -;
 DR Genew; HGNC:3761; FLRT2.
 DR MIM; 604807; -;
 DR InterPro; IPR003961; FN_III.
 DR InterPro; IPR001611; LRR.
 DR InterPro; IPR000483; LRR_Cterm.
 DR InterPro; IPR000372; LRR_Nterm.
 DR InterPro; IPR003592; LRR_Out.
 DR InterPro; IPR003591; LRR_Typ.
 DR Pfam; PF00041; Fr3; 1.
 DR Pfam; PF00560; LRR; 8.
 DR Pfam; PF01462; LRRNT; 1.
 DR Pfam; PF01463; LRRCT; 1.
 DR SMART; SM00370; LRR; 3.
 DR SMART; SM00082; LRRCT; 1.
 DR SMART; SM00013; LRRNT; 1.
 DR SMART; SM00369; LRR_Typ; 1.
 KW Transmembrane; Leucine-rich repeat; Repeat; Glycoprotein; Signal.
 FT SIGNAL 1 35
 FT CHAIN 36 660
 FT LEUCINE-RICH REPEAT TRANSMEMBRANE PROTEIN
 FT POTENTIAL.
 FT FLRT2
 FT EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 542 562
 FT DOMAIN 563 660
 FT REPEAT 62 87
 FT REPEAT 88 108
 FT REPEAT 109 131
 FT REPEAT 132 157
 FT REPEAT 159 181
 FT REPEAT 183 202
 FT REPEAT 203 228
 FT REPEAT 229 251

FT REPEAT 148 171 LRR 4.
 FT REPEAT 172 195 LRR 5.
 FT REPEAT 197 219 LRR 6.
 FT REPEAT 407 430 LRR 7.
 FT REPEAT 431 454 LRR 8.
 FT REPEAT 435 478 LRR 9.
 FT REPEAT 480 502 LRR 10.
 FT REPEAT 503 526 LRR 11.
 FT REPEAT 528 550 LRR 12.
 FT DOMAIN 712 723 POLY-GLY.
 SQ SEQUENCE 977 AA; 109005 MW; 3C936B7E0003DF54 CRC64;

Query Match 5.8%; Score 236; DB 1; Length 977;
 Best Local Similarity 24.8%; Pred. No. 1.2e-07;
 Matches 111; Conservative 42; Mismatches 178; Indels 116; Gaps 15;

QY 21 CPKYCVQNLSESLGTL-CPSKGLLFVPPDIDR--RTVELRLGGNFIIHISRODFANMT 76
 DB 373 CPTGCTCNLHINDLGLTVNCKERGFNNISELLPRPLNAKKLYLSSNLIQKIYRSDFWNFS 432
 QY 77 GLVDLTLRMTIISHIOPFSLDLESRLHLDNRLPFLSGEDTLRGLVNLQHLIVNNQL 136
 DB 433 SLDLLHLGNRRISYVQDGAFINLNLKSLFLNGNDIEKLTPTGMFRGLQSLHYLYFNFVI 492
 QY 137 GGIADAEAFEDFLTLLEDLDSYNNLHGLPWDSVRRVWNLHQLSLDHLNLDHIAEGTFADL 196
 DB 493 REIQPAAFS-----LMPNLKLLFLNNLRLTPTDAFAG- 526
 QY 197 QKLARLDLTSNRLOKLPDPPIFARSOASALTATPFAPPLSFGGNPLHCNCELL-----W 252
 DB 527 TSLARLNLRKRYFLYLPVAGVLEHLNAI-----VOIDLNENWDCTCDLVFPKQW 576
 QY 253 LRRLEER---DDDLLETCSGGKGRYFWHVR--EEFVCEPPLIQTHTKLLVLEGOAA 306
 DB 577 IETISSVYVGDVL--CESPENLTHR---DVRTIEVLVC----- 611
 QY 307 TLKCKAIGDPSPLIHWVAP-----DRLVGNSSRTAVYD---NGTLDFITTSQD 353
 DB 612 -----PEMLHVAPAGESPAQFGDLSHLIGAPTSASPYEFPSPGPPVPLSVLILSL 660
 QY 354 SGAFTCIAANAAGEATAMVEVSIVOLPHLSNSTSRTPAPPKSRSLSDITGSSKTSR----- 407
 DB 661 LVLFSSAVFVAAGLFAYVLRRRRKKLPFRSK-----RQEGVDLITGIQOCHRLEFDG 712
 QY 408 ---CGGGGGGGEPKPPPERAVLVSEV 431
 DB 713 GGGGGGGGGGRPTLSSPEKAPPVGHV 739

RESULT 13
 ALS_PAPHA STANDARD; PRT; 605 AA.
 AC O02833.
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Insulin-like growth factor binding protein complex acid labile chain precursor (ALS).
 DE IGFALS OR ALS.
 GN IGFALS hamadryas (Hamadryas baboon).
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 OC Cercopithecoidea; Papio.
 NCBI_TaxID=9557;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RX MEDLINE=97040714; PubMed=8986027;
 RA Delhanty P., Baxter R.C.;
 RT "The cloning and expression of the baboon acid-labile subunit of the insulin-like growth factor binding protein complex."
 RL Biochem. Biophys. Res. Commun. 227:897-902(1996).
 CC !- FUNCTION: INVOLVED IN PROTEIN-PROTEIN INTERACTIONS THAT RESULT

Db 13 GLLA--SLLPALAACPNQCHCH--SDLOHVICDKVGLQKI-PKVESEKTKLNLQRRNPPV 67
 QY 67 ISRODFANMTGLVD-----LTLRNTIISHIQPFSELDLES 102
 Db 68 LATSFRAMPNLSLHLOHCQIREVAAGAFRGLKQIYLYLSHNDIRVLRAGAFDDLTEL 127
 QY 103 RSLHLDNRLPSLGBDITLRLGLVNLQHLIVNNQGLGGIADAEFE----- 145
 Db 128 TYLYLDHNKVTLPGLLSPLVNLFILOLNKIKRELRSAGAFQAKDLRWLYLSENSLSS 187
 QY 146 -----DFLLTLEDLDSYNNLHGLPWDSVRRMVN- 174
 Db 188 LQPGALDDVENLAKYLDNRNQLSSVPSAALSRLRVEELKLSHPDKIPDNARQSPGRY 247
 QY 175 LQLSLDHLNLDHIAEGTFADLQKLRDLTSNRLOKLPDPPIFARSOASALTATPFAPP 234
 Db 248 LETLWLDNTNLEKFSGDGAFGLVTTLKHVHLENRLHQLPSN--FPFDSLETTLTLT----- 300
 QY 235 LSFSGGNPLHCNCELLMRR-----LERDDLETCSGGKGRYFWHVRREE 283
 Db 301 -----NPNWPKCTQLRGLRRLWEAKTSRPP--ATCASPAKFRGQ---HIRDTD 343

RESULT 12
 Y848 HUMAN STANDARD; PRT; 977 AA.
 AC O94933;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical protein KIAA0848.
 GN KIAA0848.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=99156230; PubMed=10048485;
 RA Nagase T., Ishikawa K.-I., Suyama M., Kikuno R., Hirose M., Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;
 RT "Prediction of the coding sequences of unidentified human genes. XII. The complete sequences of 100 new cDNA clones from brain which code for large proteins in vitro.;
 RL DNA Res. 5:355-364(1998).
 CC !- SIMILARITY: CONTAINS 12 LEUCINE-RICH REPEATS (LRR).
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sb-sib.ch).
 CC
 CC EMBL; AB020655; BAA74871.1;
 DR InterPro; IPR001611; LRR.
 DR InterPro; IPR000483; LRR_Cterm.
 DR InterPro; IPR000372; LRR_Nterm.
 DR InterPro; IPR003592; LRR.Out.
 DR InterPro; IPR003591; LRR_Typ.
 DR Pfam; PF00560; LRR; 9.
 DR Pfam; PF01463; LRRCT; 2.
 DR SMART; SM00370; LRR; 4.
 DR SMART; SM00082; LRRCT; 2.
 DR SMART; SM00013; LRRNT; 1.
 DR SMART; SM00369; LRR_Typ; 3.
 KW Hypothetical protein; Repeat; Leucine-rich repeat.
 FT REPEAT 76 99 LRR 1.
 FT REPEAT 100 123 LRR 2.
 FT REPEAT 125 147 LRR 3.

Repeat; Leucine-rich repeat. POTENTIAL.
 KW SIGNAL 1 24
 FT CHAIN 25 951
 FT FT LEUCINE-RICH REPEAT-CONTAINING G PROTEIN-
 FT FT COUPLED RECEPTOR 4.
 FT FT EXTRACELLULAR (POTENTIAL).
 FT FT 1 (POTENTIAL).
 FT FT 2 (POTENTIAL).
 FT FT 3 (POTENTIAL).
 FT FT 4 (POTENTIAL).
 FT FT 5 (POTENTIAL).
 FT FT 6 (POTENTIAL).
 FT FT 7 (POTENTIAL).
 FT FT 8 (POTENTIAL).
 FT FT 9 (POTENTIAL).
 FT FT 10 (POTENTIAL).
 FT FT 11 (POTENTIAL).
 FT FT 12 (POTENTIAL).
 FT FT 13 (POTENTIAL).
 FT FT 14 (POTENTIAL).
 FT FT 15 (POTENTIAL).
 FT FT BY SIMILARITY.
 FT FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT FT F -> S (IN REF. 1; AA#68989).
 FT FT L -> P (IN REF. 1; AA#68989).
 FT FT L -> S (IN REF. 1; AA#68989).
 FT FT SEQUENCE 951 AA; 104460 MW; 5E0C2DFCF22CA1BB CRC64;
 Query Match 5.7%; Score 234.5; DB 1; Length 951;
 Best Local Similarity 22.0%; Pred. No. 1.4e-07;
 Matches 148; Conservative 79; Mismatches 251; Indels .195; Gaps 27;
 QY 7 GLLAF-----GMAFAVADCPKVCYQNLSESLGTLCPKGLLFPVPPIDRRTVEL 57
 Db 6 GLLCFALGLLGSAGPSGAAPLCAAPCSDG---DRRVDCSGKGLTAVPEGLSFTQAL 62
 QY 58 RLGGNFIHHSRODFANMGLVDLTSRNTISHIQPFSLDLE----- 100
 Db 63 DISMNNITQLPEDAFKNFPFLEQLAGNDLSFIHPKALSGLKLKLVLLQNNLKVPS 122
 QY 101 -----SLRSRLHSDNPLSGEDTLRGLVNLQHLVNNNOLGGTAEAFEDFLTLLED 153
 Db 123 EAIRGUSALQSLDANHITSVPEDSFEGLVQRHLWLDNLSLTVVPHPLSN-LPTLQA 181
 QY 154 LDLSYNNLHGLPMSVRRMVNHLQSLDNLHLDHIAEGTFADLQKRLARLDLTSNRLOKLP 213
 Db 182 LTLALNKISSIPDAFTNLSLVLVHLHNNKIRGLSQHCFGLDNLDTLDSYNNLGEFP 241
 QY 214 PDPFARSQASALTATPPAPLPSF-----SFGGNPLHCNCELLWRLERDDDL 262
 Db 242 -----QAIKARPSLKEGLGFHSNSISVDPGAFDGNPL-----LRTHLYD-- 281
 QY 263 ETCGSPGGLKGRFYFH-----VREBEFYCEPPLITQHTH-KLLVLEGOAATLKCKA 312
 Db 282 ----NPLSFGNSAFHNSLDHSLVIRGASMQQFPNLTGTVHLESLLTG-----YKI 331

QY 313 IGDPSPLIHWVADPDRRLVGNSSRTAVYDNGTLDIFITTSQDSGAPT-CIAAN----- 363
 Db 332 SSIENNL-----COEQKML-----RTLDLSYNNIRDLPSFENGCHALEEISLQRNQ 376
 QY 364 ---AAGEATAMVEYSIVQLP-----HLSNSTSRTPAPKSRLLSDITGSSKTSRGGGGSGG 414
 Db 377 IYQIKEGFQGLISLRILDLRNLRIHEIHSRAFAFLGPIITNL-DVSFNLHLSFFPTEGLNG 435
 QY 415 GEPKPSPPERAVLVSEVTTTSALVKWSVSKSAPRVKMYQLQVNCSDDEVLVIRMI PASNK 474
 Db 436 LNQLK-----LVGNFKLKEAL-----AAKDFVNLRSLSVPY----- 466
 QY 475 AFVNNLVSGTYDLCLVLA MW--DDTATLTLATNIV---GCAQFTFKADYPOCQSM---- 525
 Db 467 -----AYOCC--AFWGCDSYANLNTEDNSLODHSVAQERGTADAANVTSTLENE 513
 QY 526 -HSQI-----LGGTMI-LVIGGIIVATLLVFIIVILVMRYKVCNHEAPS 566
 Db 514 EHSQIIHCTPSTGAFKPCYLLGSMIRLTVWFIPLVALFFNLVLLVLTTFASCTSLPSS 573
 QY 567 KM--AAAVSNVY 576
 Db 574 KLFGLISVSNLF 586
 ID_LGR7 HUMAN STANDARD; PRT; 757 AA.
 AC Q9HEX9.
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Relaxin receptor 1 (Leucine-rich repeat-containing G protein-coupled receptor 7).
 GN LGR7.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N. A. (ISOFORMS 1 AND 2), AND MUTAGENESIS OF ASP-637.
 RX MEDLINE=20388592; PubMed=10935549;
 RA Hsu S.Y., Kudo M., Chen T., Nakabayashi K., Bhalla A.,
 RA van der Spek P.J., Van Duin M., Hsueh A.J.W.;
 RT "The three subfamilies of leucine-rich repeat-containing G protein-
 RT coupled receptors (LGR): identification of LGR6 and LGR7 and the
 RT signaling mechanism for LGR7";
 RL Mol. Endocrinol. 14:1257-1271(2000).
 RN [2]
 RP CHARACTERIZATION.
 RP PubMed=11517286;
 RA Bartsch O., Bartlick B., Ivell R.;
 RT "Relaxin signalling links tyrosine phosphorylation to
 RT phosphodiesterase and adenylyl cyclase activity.";
 RL Mol. Hum. Reprod. 7:799-809(2001).
 CC -!- FUNCTION: Receptor for relaxin. The activity of this receptor is
 CC mediated by G proteins leading to stimulation of adenylyl cyclase
 CC and an increase of cAMP. Binding of the ligand may also
 CC activate a tyrosine kinase pathway that inhibits the activity of a
 CC phosphodiesterase that degrades cAMP.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -!- ALTERNATIVE PRODUCTS: 2 isoforms; 1 (shown here) and 2; are
 CC produced by alternative splicing.
 CC -!- TISSUE SPECIFICITY: Expressed in the brain, kidney, testis,
 CC placenta, uterus, ovary, adrenal, prostate, skin and heart. Not
 CC detected in spleen.
 CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC -!- SIMILARITY: CONTAINS 1 LDL-RECEPTOR CLASS A DOMAIN.
 CC -!- SIMILARITY: CONTAINS 10 LEUCINE-RICH REPEATS (LRR).
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its

```

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.iab-sib.ch/announce/
CC or send an email to license@iab-sib.ch)
CC -----
DR EMBL; AF190500; AAG17167.1; .
DR MIM; 606854; .
DR HSP; P01130; ILDR.
DR InterPro; IPR000276; GPCR_rhoqpsn.
DR InterPro; IPR002172; LDL_recept_A.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR000372; LRR_Nterm.
DR InterPro; IPR003592; LRR_out.
DR InterPro; IPR003591; LRR_typ.
DR Pfam; PF00001; 7tm_1; 1.
DR Pfam; PF00057; ldl_recept_a; 1.
DR Pfam; PF00560; LRR; 8.
DR PRINTS; PF00237; GPCRHOQPSN.
DR SMART; SM00192; LDL_e; 1.
DR SMART; SM00370; LRR; 7.
DR SMART; SM00013; LRRNT; 1.
DR SMART; SM00369; LRR_TYP; 10.
DR PROSITE; PS00237; G_PROTEIN_RECP_F1_1; FALSE_NEG.
DR PROSITE; PS0262; G_PROTEIN_RECP_F1_2; 1.
DR PROSITE; PS01209; LDLRA_1; 1.
DR PROSITE; PS00068; LDLRA_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein; Repeat;
KW Leucine-rich repeat; Alternative splicing.
FT DOMAIN 1 409
FT TRANSMEM 410 430
FT DOMAIN 431 443
FT TRANSMEM 444 464
FT DOMAIN 465 486
FT TRANSMEM 487 507
FT DOMAIN 508 527
FT TRANSMEM 528 548
FT DOMAIN 549 577
FT TRANSMEM 578 598
FT DOMAIN 599 629
FT TRANSMEM 630 650
FT DOMAIN 651 651
FT TRANSMEM 652 672
FT DOMAIN 673 757
FT DOMAIN 26 63
FT REPEAT 125 148
FT REPEAT 150 172
FT REPEAT 173 196
FT REPEAT 198 220
FT REPEAT 221 244
FT REPEAT 246 269
FT REPEAT 270 293
FT REPEAT 294 317
FT REPEAT 319 341
FT REPEAT 343 365
FT DISULFID 485 563
FT CARBOHYD 36 36
FT CARBOHYD 127 127
FT CARBOHYD 264 264
FT CARBOHYD 272 272
FT CARBOHYD 325 325
FT CARBOHYD 368 368
FT VARSPLIC 63 96
FT MUTAGEN 637 637
SQ SEQUENCE 757 AA; 86992 MW; 8079E8DF3A3EF21 CRC64;

```

```

Query Match      5.7%; Score 232.5; DB 1; Length 757;
Best Local Similarity 31.8%; Pred. No. 1.4e-07;
Matches 69; Conservative 34; Mismatches 95; Indels 19; Gaps 3;

```

```

Query 17 VVDACPKYVCQNLS-----ESLGLTLPFGKGLLF-----VPPDIRRVLELR--58
| : | | | | | | | | | | | : | | | | | | : | | | | | |
: : | | | | | | | | | | | : | | | | | | : | | | | | |

```

```

96 LVGSVPVQCLCOGLELDCCDETNLRVPVSSVTAMSLQWNLIRKLPDPDCKVYHDLOKL 155
59 -LGGNFIIHISRQDFANMTGLVDITLSRNTIISHIQPFSLDLLESLSLHDSNRPLPSLGE 117
156 YLQNNKITSISIVAFRGLNSLTLYLSHNRITLTKVGFVDFDLHRLEWLIIEDNHLRSRISP 215
118 DTLRGLVNLQHLIWNVNNQLGGIADAEAFEDFLLEDDLDLSYNNLHGLPWDSVRRMVLHQ 177
216 PTFYGLNSLLVLMNNVLRPLPKPLCQHMPLRLHLWLDLEGNHIHNLRLNTFFIYCSNLTV 275
178 LSLDHNLLDHIAGCTPADLQKARLDLTSNRKLOKLPP 214
276 LVWRKWKINHLNENTFAPLQKLDLDELGSGNKNIEIENLPP 312

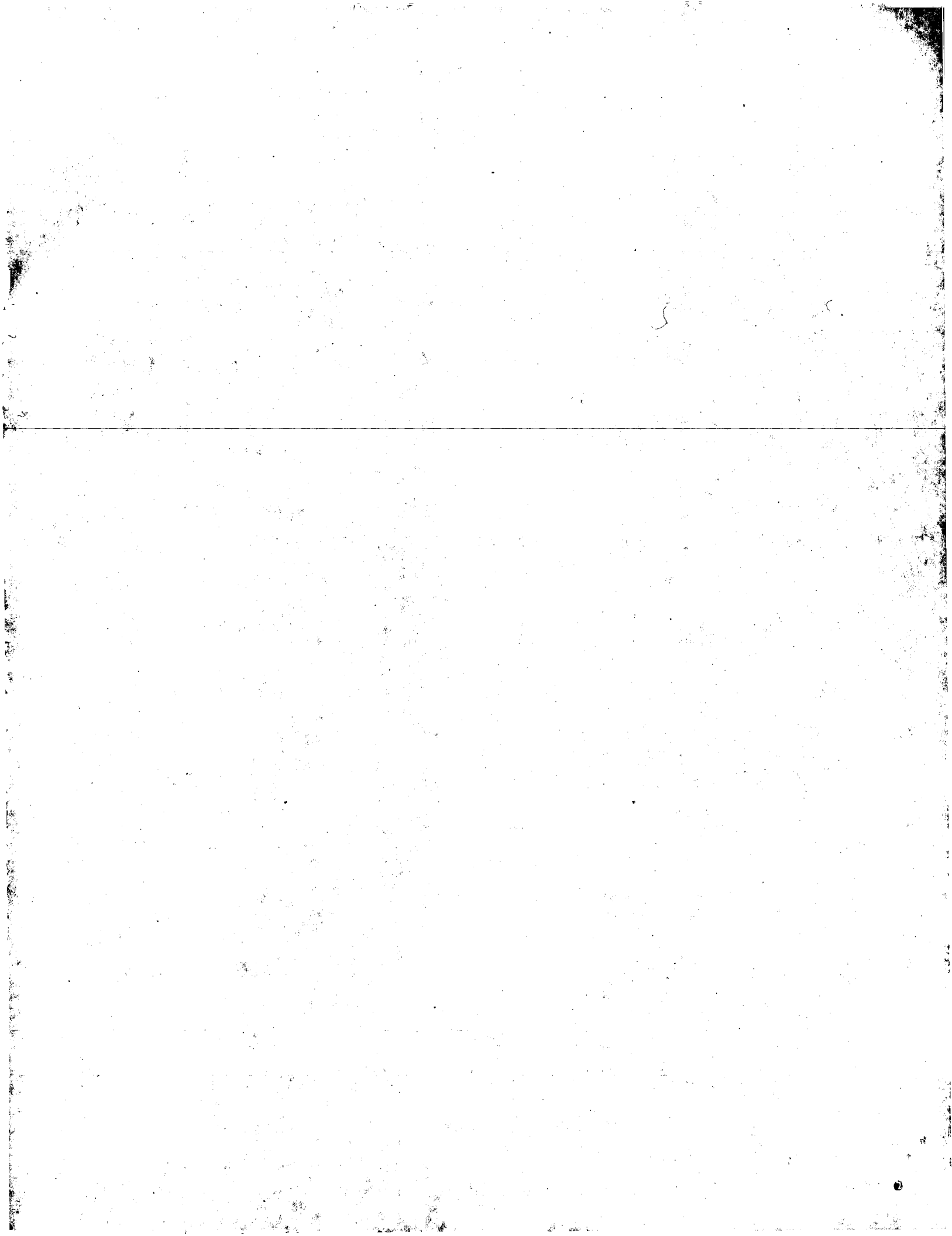
```

Search completed: June 8, 2003, 21:29:07
Job time : 41 secs

```

Query 17 VVDACPKYVCQNLS-----ESLGLTLPFGKGLLF-----VPPDIRRVLELR--58
| : | | | | | | | | | | | : | | | | | | : | | | | | |
: : | | | | | | | | | | | : | | | | | | : | | | | | |

```



GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 8, 2003, 20:50:18 ; Search time 62 Seconds
(without alignments)
2622.116 Million cell updates/sec

Title: US-09-831-846-2
Perfect score: 4094
Sequence: 1 METLLGGLLAFGMFAVVDA.....DLVGARGTFGSSEWVMESTV 789

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SPTREMBL_21.*

- 1: sp_arches.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phage.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_rvirus.*
- 16: sp_bacteriaip.*
- 17: sp_archaeip.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	4094	100.0	832	4 Q9ULH4	Q9ulh4 homo sapien
2	4075	99.5	789	6 Q9BE71	Q9be71 macaca fasc
3	3885.5	94.9	788	11 Q9CYK3	Q9cyk3 mus musculus
4	1821	44.5	719	4 Q9GN16	Q9gn16 homo sapien
5	1573	38.4	628	4 Q9BTN0	Q9btn0 homo sapien
6	1571.5	38.4	700	4 Q9P244	Q9p244 homo sapien
7	1320.5	32.3	492	11 Q99KT6	Q99kt6 mus musculus
8	1063.5	26.0	542	4 Q8TBS9	Q8tbs9 homo sapien
9	775	18.9	450	4 Q96C50	Q96c50 homo sapien
10	653	16.0	324	4 Q9BQJ0	Q9bqj0 homo sapien
11	435.5	10.6	640	4 Q9HCJ2	Q9hcj2 homo sapien
12	418.5	10.2	606	4 Q9BZ20	Q9bz20 homo sapien
13	385	9.4	649	4 Q96A85	Q96a85 homo sapien
14	383.5	9.4	648	11 Q9VIJ5	Q9vij5 mus musculus
15	383	9.4	653	4 Q9HBW1	Q9hbw1 homo sapien
16	380.5	9.3	707	11 Q9ESY6	Q9esy6 rattus norv

ID	Q9ULH4	PRELIMINARY;	PRT;	832 AA.
AC	Q9ULH4;			
DT	01-MAY-2000 (TremBLrel. 13, Created)			
DT	01-MAY-2000 (TremBLrel. 13, Last sequence update)			
DT	01-DEC-2001 (TremBLrel. 19, Last annotation update)			
DE	KIAA1246 protein (Fragment).			
GN	KIAA1246.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=BRAIN;			
RX	MEDLINE=20039619; PubMed=10574462;			
RA	Nagase T., Ishikawa K., Kikuno R., Hirose M., Nomura N., Ohara O.,			
RT	"Prediction of the coding sequences of unidentified human genes. XV.			
RT	The complete sequences of 100 new cDNA clones from brain which code			
RL	DNA Res. 6:337-345 (1999)			
DR	EMBL; AB033072; BA86560.1			
DR	HSSP; P56276; 1TLK			
DR	InterPro; IPR003961; FN_III.			
DR	InterPro; IPR003598; IG_C2.			
DR	InterPro; IPR003006; IG_MHC.			
DR	InterPro; IPR001611; LRR.			
DR	InterPro; IPR000483; LRR_Cterm.			
DR	InterPro; IPR000372; LRR_Nterm.			
DR	InterPro; IPR003592; LRR_out.			
DR	InterPro; IPR003591; LRR_typ.			
DR	Pfam; PF00041; fn3; 1.			
DR	Pfam; PF00560; LRR; 7.			
DR	Pfam; PF01463; LRRCT; 1.			
DR	PRINTS; PR00019; LEURICHRPT.			
DR	SMART; SM00060; FN3; 1.			
DR	SMART; SM000408; IGC2; 1.			

ALIGNMENTS

RESULT 1

Q9ULH4 Q9ULH4 PRELIMINARY; PRT; 832 AA.

DR SMART; SM00370; LRR; 1.
 DR SMART; SM00082; LRRCT; 1.
 DR SMART; SM00013; LRRNT; 1.
 DR SMART; SM00369; LRR_TYP; 3.
 KW Immunoglobulin domain.
 FT NON_TER 1
 SQ SEQUENCE 832 AA; 88978 MW; 9B0E80F7BFF974A8 CRC64;

Query Match 100.0%; Score 4094; DB 4; Length 832;
 Best Local Similarity 100.0%; Pred. No. 5, 4e-310;
 Matches 789; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 METLGGLLAFGMFAVVDACPKYCVQNLSLGLTLCPSKGLLFPVPPIDRRRTVELRLG 60
 Db 44 METLGGLLAFGMFAVVDACPKYCVQNLSLGLTLCPSKGLLFPVPPIDRRRTVELRLG 103
 QY 61 GNFIHISRODFANMTGLVDLTLRSNTISHIQPFSLDLESLSLHLSNRLPSLGEDTL 120
 Db 104 GNFIHISRODFANMTGLVDLTLRSNTISHIQPFSLDLESLSLHLSNRLPSLGEDTL 163
 QY 121 RGLVNLQHLI VNNNOLGGIADAEFELLTLEDLDSYNNLHGLPWDSVRRMVLHQLSL 180
 Db 164 RGLVNLQHLI VNNNOLGGIADAEFELLTLEDLDSYNNLHGLPWDSVRRMVLHQLSL 223
 QY 181 DHNLDHIAEGTFADLQKLARLDLTSNRLOKLPDPPIFARSOASALTATPPAPPLSFSGF 240
 Db 224 DHNLDHIAEGTFADLQKLARLDLTSNRLOKLPDPPIFARSOASALTATPPAPPLSFSGF 283
 QY 241 GNPLHCNCELLWRLERDDLETCGSPGGLKGRYFHWVREBEFVCEPPLIQTHTKLLV 300
 Db 284 GNPLHCNCELLWRLERDDLETCGSPGGLKGRYFHWVREBEFVCEPPLIQTHTKLLV 343
 QY 301 LEGQAATLKCKAIGDPSLIHWVAPDDRLVGNSSRTAVYDNGTLDLFIITSDSQAFTCI 360
 Db 344 LEGQAATLKCKAIGDPSLIHWVAPDDRLVGNSSRTAVYDNGTLDLFIITSDSQAFTCI 403
 QY 361 AANAAGEATAMVEVSIVQLPHLSNSTRTAPPKRSLSIDITGSSKTSRGGSGGGGEPKPS 420
 Db 404 AANAAGEATAMVEVSIVQLPHLSNSTRTAPPKRSLSIDITGSSKTSRGGSGGGGEPKPS 463
 QY 421 PPERAVLVSEVTTTSALVKWSYKSAAPRVKMYOLOYNCSDDDEVLIYRMTIPASNKAPVNN 480
 Db 464 PPERAVLVSEVTTTSALVKWSYKSAAPRVKMYOLOYNCSDDDEVLIYRMTIPASNKAPVNN 523
 QY 481 LVSGTGYDLCVLAEMDDTATTATNIVGCAOFFTKADYPOCOSMHSQILGTMILVIGG 540
 Db 524 LVSGTGYDLCVLAEMDDTATTATNIVGCAOFFTKADYPOCOSMHSQILGTMILVIGG 583
 QY 541 IIVATLLVFIILMVKYKCNHEAPSKMAAANVSNVYQNGAQPPTPPSSAPAGAPQGGP 600
 Db 584 IIVATLLVFIILMVKYKCNHEAPSKMAAANVSNVYQNGAQPPTPPSSAPAGAPQGGP 643
 QY 601 KVVVRNELLDFATSLARASDSSSSLSGGEAGLGRAPWRIPPSAPRKPDSIDRLMGAF 660
 Db 644 KVVVRNELLDFATSLARASDSSSSLSGGEAGLGRAPWRIPPSAPRKPDSIDRLMGAF 703
 QY 661 ASLDLKSQRKEILLDSRTPAGRGAGTSARGHSDREPLGPPAARARSLLPLPLEGKAKR 720
 Db 704 ASLDLKSQRKEILLDSRTPAGRGAGTSARGHSDREPLGPPAARARSLLPLPLEGKAKR 763
 QY 721 SHSFDMGDFAAAAAGVPGGYSPRKRNSNIWTKRSLSVNGMLLPPFEESDLVGARTFGS 780
 Db 764 SHSFDMGDFAAAAAGVPGGYSPRKRNSNIWTKRSLSVNGMLLPPFEESDLVGARTFGS 823
 QY 781 SEMVMESTV 789
 Db 824 SEMVMESTV 832

DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE Hypothetical 84.7 kDa protein.
 OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 OC Cercopithecoidea; Macaca.
 OX NCBI_TaxID=9541;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=FRONTAL LOBE LEFT;
 RA Osada N., Hida M., Kusuda J., Tanuma R., Iseki K., Hirai M., Terao K.,
 RA Suzuki Y., Sugano S., Hashimoto K.;
 RT "Isolation of full-length cDNA clones from macaque brain cDNA
 RT libraries";
 RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB056799; BAB39323.1; -.
 DR HSSP; P56276; IILK.
 DR InterPro; IPR003961; FN_III.
 DR InterPro; IPR003599; IG.
 DR InterPro; IPR003598; IG_C2.
 DR InterPro; IPR003006; IG_MHC.
 DR InterPro; IPR001611; LRR.
 DR InterPro; IPR000483; LRR_Cterm.
 DR InterPro; IPR000372; LRR_Nterm.
 DR InterPro; IPR003592; LRR_out.
 DR InterPro; IPR003591; LRR_typ.
 DR Pfam; PF00041; fn3_1.
 DR Pfam; PF00047; Ig_1.
 DR Pfam; PF00560; LRR; 7.
 DR Pfam; PF01463; LRRCT; 1.
 DR PRINTS; PR00019; LEURICHRPT.
 DR SMART; SM00060; FN3_1.
 DR SMART; SM00409; IG_1.
 DR SMART; SM00408; IGC2_1.
 DR SMART; SM00370; LRR; 5.
 DR SMART; SM00082; LRRCT; 1.
 DR SMART; SM00013; LRRNT; 1.
 DR SMART; SM00369; LRR_TYP; 6.
 KW Hypothetical protein; Immunoglobulin domain.
 SQ SEQUENCE 789 AA; 84730 MW; BB86DE81BC284B23 CRC64;

Query Match 99.5%; Score 4075; DB 6; Length 789;
 Best Local Similarity 99.4%; Pred. No. 1.5e-308;
 Matches 784; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 METLGGLLAFGMFAVVDACPKYCVQNLSLGLTLCPSKGLLFPVPPIDRRRTVELRLG 60
 Db 1 METLGGLLAFGMFAVVDACPKYCVQNLSLGLTLCPSKGLLFPVPPIDRRRTVELRLG 60
 QY 61 GNFIHISRODFANMTGLVDLTLRSNTISHIQPFSLDLESLSLHLSNRLPSLGEDTL 120
 Db 61 GNFIHISRODFANMTGLVDLTLRSNTISHIQPFSLDLESLSLHLSNRLPSLGEDTL 120
 QY 121 RGLVNLQHLI VNNNOLGGIADAEFELLTLEDLDSYNNLHGLPWDSVRRMVLHQLSL 180
 Db 121 RGLVNLQHLI VNNNOLGGIADAEFELLTLEDLDSYNNLHGLPWDSVRRMVLHQLSL 180
 QY 181 DHNLDHIAEGTFADLQKLARLDLTSNRLOKLPDPPIFARSOASALTATPPAPPLSFSGF 240
 Db 181 DHNLDHIAEGTFADLQKLARLDLTSNRLOKLPDPPIFARSOASALTATPPAPPLSFSGF 240
 QY 241 GNPLHCNCELLWRLERDDLETCGSPGGLKGRYFHWVREBEFVCEPPLIQTHTKLLV 300
 Db 241 GNPLHCNCELLWRLERDDLETCGSPGGLKGRYFHWVREBEFVCEPPLIQTHTKLLV 300
 QY 301 LEGQAATLKCKAIGDPSLIHWVAPDDRLVGNSSRTAVYDNGTLDLFIITSDSQAFTCI 360
 Db 301 LEGQAATLKCKAIGDPSLIHWVAPDDRLVGNSSRTAVYDNGTLDLFIITSDSQAFTCI 360
 QY 361 AANAAGEATAMVEVSIVQLPHLSNSTRTAPPKRSLSIDITGSSKTSRGGSGGGGEPKPS 420

RESULT 2
 Q9BE71 PRELIMINARY; PRT; 789 AA.
 ID Q9BE71
 AG Q9BE71;

Db 361 ANAAGEATATVEVSIQVLPVPHLSNSTSRTPPKSRSLSDITGSKTSRGGGGGGGEPKPKS 420
 QY 421 PPERAVLVEVTTTTSALVWVSKSAPRYKMYQVNCSDDEVLVYRMIIPASNKAFVNN 480
 Db 421 PPERAVLVEVTTTTSALAKWSKSTPRVEMVQVNCSDDEVLVYRMIIPASNKAFVNN 480
 QY 481 LVSTGYDLCVLAWDDTATTLTATNIVGCAQFFTKADYPOCOSMHSQILGTMILVIGG 540
 Db 481 LVSTGYDLCVLAWDDTATTLTATNIVGCAQFFTKADYPOCOSMHSQILGTMILVIGG 540
 QY 541 IIVATLLVFIIVLMVRYKVCNHEARSKMAAAVSNVYQVNTGAPPPSSAPAGAPQGGP 600
 Db 541 IIVATLLVFIIVLMVRYKVCNHEARSKMAAAVSNVYQVNTGAPPPSSAPAGAPQGGP 600
 QY 601 KVVVRELLDFTASLARASDSSSSSLGSGEAGLGRAPWRIPPSAPRPKPSLDRLMGAF 660
 Db 601 KVVVRELLDFTASLARASDSSSSSLGSGEAGLGRAPWRIPPSAPRPKPSLDRLMGAF 660
 QY 661 ASLDLKSQRKEELLDSRTPAGRGAGTSARGHSDREPLILGPPAARARSLLPLLEGKAKR 720
 Db 661 ASLDLKSQRKEELLDSRTPAGRGAGTSARGHSDREPLILGPPAARARSLLPLLEGKAKR 720
 QY 721 SHSFDMPGFAAAAAGVPGGYSPPRKNVNIWKSLVNGMLLPEEESDLVGARTFGS 780
 Db 721 SHSFDMPGFAAAAAGVPGGYSPPRKNVNIWKSLVNGMLLPEEESDLVGARTFGS 780
 QY 781 SEWMVESTV 789
 Db 781 SEWMVESTV 789

RESULT 3
 Q9CYK3
 ID Q9CYK3 PRELIMINARY; PRT: 788 AA.
 AC Q9CYK3
 DT 01-JUN-2001 (TReMBLrel. 17, Created)
 DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
 DE 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
 DE 5730420005RIK protein.
 GN 5730420005RIK.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=EMBRYO;
 EX MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleschmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
 RA Kuelh P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
 RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
 RA Hayashizaki Y.;
 RT "Functional annotation of a full-length mouse cDNA collection."
 RL Nature 409:685-690(2001).
 DR EMBL; AK017594; BAB30828.1; -.
 DR HSSP; P56276; 1TLK.
 DR MGD; MGI:1917780; 5730420005RIK.
 DR InterPro; IPR003961; FN.III.
 DR InterPro; IPR003599; IG.

DR InterPro; IPR003598; IG_C2.
 DR InterPro; IPR003006; IG_MHC.
 DR InterPro; IPR001611; LRR.
 DR InterPro; IPR000483; LRR_Nterm.
 DR InterPro; IPR000372; LRR_Cterm.
 DR InterPro; IPR003591; LRR_typ.
 DR Pfam; PF00041; fn3; 1.
 DR Pfam; PF00047; IG; 1.
 DR Pfam; PF00560; LRR; 7.
 DR Pfam; PF01463; LRRCT; 1.
 DR PRINTS; PRO0019; LEURICHRPT.
 DR SMART; SM00060; FN3; 1.
 DR SMART; SM00409; IG; 1.
 DR SMART; SM00408; IGC2; 1.
 DR SMART; SM00082; LRRCT; 1.
 DR SMART; SM00013; LRRNT; 1.
 DR SMART; SM00369; LRR_TYP; 6.
 KW Immunoglobulin domain
 SQ SEQUENCE 788 AA; 85076 MW; 70B70C5E5B52277B CRC64;
 Query Match 94.9%; Score 3885.5; DB 11; Length 788;
 Best Local Similarity 94.9%; Pred. No. 8.6e-294;
 Matches 750; Conservative 10; Mismatches 27; Indels 3; Gaps 2;
 QY 1 METLLGGLLAFGMFAVVDACPKYCVQNLSESLGTLCPKGLLFPVPPIDRRRTVELRLG 60
 Db 1 METLLGGLLAFGMFAVVDACPKYCVQNLSESLGTLCPKGLLFPVPPIDRRRTVELRLG 60
 QY 61 GNFIHIGRODPANMTGLVDTLSRNTLSHQPFSFLDLESURSLHLSNRLPSLGEDTL 120
 Db 61 GNFIHIGRODPANMTGLVDTLSRNTLSHQPFSFLDLESURSLHLSNRLPSLGEDTL 120
 QY 121 RGLVNLQHLIYNNNQLGGIADEAFEDFLTLLEDDLNSYNNLHGLPWSVRRMVLHQLSL 180
 Db 121 RGLVNLQHLIYNNNQLGGIADDAFEDFLTLLEDDLNSYNNLHGLPWSVRRMVLHQLSL 180
 QY 181 DNLHLDHIAEGTFADLQKLARLDLTSNRLOKLPDPPIFARSOASALTATPPAPPLSFSFG 240
 Db 181 DNLHLDHIAEGTFADLQKLARLDLTSNRLOKLPDPPIFARSOASALTATPPAPPLSFSFG 240
 QY 241 GNPLHCNCELLWLRLELREDDLETCGSPGGKGRYFVHVRBEFVCEPPLITQHTKLLV 300
 Db 241 GNPLHCNCELLWLRLELREDDLETCGSPGGKGRYFVHVRBEFVCEPPLITQHTKLLV 300
 QY 301 LEGQAATLKCAIGDPPSLIHWVAPDDRLVGNSSRTAVYDNGTLDIFITTSQDSGAFTCI 360
 Db 301 LEGQAATLKCAIGDPPSLIHWVAPDDRLVGNSSRTAVYDNGTLDIFITTSQDSGAFTCI 360
 QY 361 ANAAGEATATVEVSIQVLPVPHLSNSTSRTPPKSRSLSDITGSKTSRGGGGGGGEPKPKS 420
 Db 361 ANAAGEATATVEVSIQVLPVPHLSNSTSRTPPKSRSLSDITGSKTSRGGGGGGGEPKPKS 420
 QY 421 PPERAVLVEVTTTTSALVWVSKSAPRYKMYQVNCSDDEVLVYRMIIPASNKAFVNN 480
 Db 421 PPERAVLVEVTTTTSALVWVSKSAPRYKMYQVNCSDDEVLVYRMIIPASNKAFVNN 480
 QY 481 LVSTGYDLCVLAWDDTATTLTATNIVGCAQFFTKADYPOCOSMHSQILGTMILVIGG 540
 Db 481 LVSTGYDLCVLAWDDTATTLTATNIVGCAQFFTKADYPOCOSMHSQILGTMILVIGG 540
 QY 541 IIVATLLVFIIVLMVRYKVCNHEARSKMAAAVSNVYQVNTGAPPPSSAPAGAPQGGP 599
 Db 541 IIVATLLVFIIVLMVRYKVCNHEARSKMAAAVSNVYQVNTGAPPPSSAPAGAPQGGP 599
 QY 600 KVVVRELLDFTASLARASDSSSSSLGSGEAGLGRAPWRIPPSAPRPKPSLDRLMGAF 659
 Db 600 KVVVRELLDFTASLARASDSSSSSLGSGEAGLGRAPWRIPPSAPRPKPSLDRLMGAF 659
 QY 660 ASLDLKSQRKEELLDSRTPAGRGAGTSARGHSDREPLILGPPAARARSLLPLLEGKAKR 719
 Db 660 ASLDLKSQRKEELLDSRTPAGRGAGTSARGHSDREPLILGPPAARARSLLPLLEGKAKR 719
 QY 720 SHSFDMPGFAAAAAGVPGGYSPPRKNVNIWKSLVNGMLLPEEESDLVGARTFG 779

```

Db 721 RGHSEFDMGDAFAAAA--AVPGGYSPPRRVSNWTKRSLVNGMLLFPFESDLVGARGTFG 778
QY 780 SSEWMESTV 789
Db 779 SSEWMESTV 788

RESULT 4
ID Q96NI6 PRELIMINARY; PRT; 719 AA.
AC Q96NI6;
DT 01-DEC-2001 (TRENBLrel. 19, Created)
DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)
DT 01-JUN-2001 (TRENBLrel. 21, Last annotation update)
DE CDNA FLJ30803 fis, Clone FEBRA2001245, weakly similar to NAG14.
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN;
RA Ninomiya K., Wagatsuma M., Kanda K., Kondo H., Yokoi T., Kodaira H.,
RA Furiya T., Takahashi M., Kikkawa E., Omura Y., Abe K., Kamihara K.,
RA Katsuta N., Sato K., Tanikawa M., Yamazaki M., Sugiyama T., Irie R.,
RA Otsuki T., Sato H., Wakamatsu A., Ishii S., Yamamoto J., Isono Y.,
RA Kawai-Hio Y., Saito K., Nishikawa T., Kimura K., Yamashita H.,
RA Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Murakawa K.,
RA Kanehori K., Takahashi-Fujii A., Oshima A., Sugiyama A., Kawakami B.,
RA Suzuki Y., Sugano S., Nagahari K., Maeuho Y., Nagai K., Isogai T.;
RT "NEO human cDNA sequencing project."
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK055365; BAB70910.1;
DR InterPro; IPR003961; FN_III.
DR InterPro; IPR003006; IG_MHC.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR000483; LRR_Cterm.
DR InterPro; IPR001005; Myb_DNA_binding.
DR Pfam; PF00041; fn3; 1.
DR Pfam; PF00047; ig; 1.
DR Pfam; PF00560; LRR; 6.
DR Pfam; PF01463; LRRCT; 1.
DR PROSITE; PS00037; MYB_1; UNKNOWN 1.
SQ SEQUENCE 719 AA; 79472 MW; CE301F219705FF4B CRC64;

Query Match 44.5%; Score 1821; DB 4; Length 719;
Best Local Similarity 51.4%; Pred. No. 4.6e-133;
Matches 375; Conservative 118; Mismatches 195; Indels 42; Gaps 13;

QY 1 METLGLLAFGMAFAVAVDACPKYCVQNLSESLGTLCPKGLLFPVPPIDRRITVELRUG 60
Db 1 MEKILEYFLIGIA-VKAIQCPKRCVQLSPNLATLCAKGLLFFVPPNIDRRITVELRLA 59

QY 61 GNFIHISRQDFANMTGLVDLTLNRTISHIOPFSELDLESLSRLSLDNLPSLGEDTL 120
Db 60 DNFVTVIKRDFANMTSLVDLTLNRTISFIPTPHAFADLRNLRALHLSNRLTKITNDF 119

QY 121 RGLVNLQHLIVANNLGGIADAEFELITLEDLDLSYNNLHGLPWSVRRVNLHQLSL 180
Db 120 SGLSNLHLILNANNQLTSLSTAFDD-VFALBELDLSYNNLETIPMDAVERKWSLHTLSL 178

QY 181 DNLNLDHIAEGTFADLQKLARDLTSNRLQKLPDPPIFARSOASALATATPFPAPLPSFSG 240
Db 179 DNMWIDNIKGTFSHLHKTRLDVTSNKLQKLPDPDLFORAQLVATSGIISSTALSFG 238

QY 241 GNPLHCNCELLMRLRERDDLETCCSPGGLKGRYFHWVREBEFVCEPPLITQHTHLIV 300
Db 239 GNPLHCNCELLMRLRERDDLETCCSPGGLKGRYFHWVREBEFVCEPPLITRHTHEMRV 298

QY 301 LEGOATLCKAIGDPSPLIHWVAPDRLVGNSSRAVDNGLTDLFFITTSQDSGAFTCI 360
Db 299 LEGOATLCKARGDPEPAIHWISPEGLKISNATRSVLAVDNGTDLILITTVKDTGAFTCI 358

```

```

QY 361 AANAAGEATAMVEYSIVQLPHLSNSTSTAPPKRLSLDITGSSKTSRGGGSGGGEPPKKS 420
Db 359 ASNPAGEATQVLDLHIKPLHLLNSTNHIHEPDPGSSDITSTKSGSNTSSNG---DTK 415

QY 421 PPERAVLVSEVTTTSALVKWSYSKSAFRVKMYQIOLYNSDDDEVLIYRMIIPASNKAFVNN 480
Db 416 LSODKIVVAEATSSALLKFNQIRNIPGIRMFQIQYNGTYDDTLVYRMIIPSTKFTLVNN 475

QY 481 LVSGTGYDLCLVAMWDDTATLTATNIVGCAQFFTKADYPOCSMHSOILGGTMLVIGG 540
Db 476 LAAGMYDLCLVALYDDGITSLTATRVVGCCLQFTTEQDYVRFCHFMQSFQGGTMIILIGG 535

QY 541 IIVATLLVFIILMRYKVCNHEAPSKMAAAVSNVYQTNGAQPPPPSPSAPAGAPPQPPP 600
Db 536 IIVASVLFVLIILMRYKVCNNGGHKV-TRVSNVYQTNGAQ----IQGCVTLTPQSVS 590

QY 601 KVVVRNELLDTASLARASDSS---SSSLSGGEAAGLGRAPWRIPPSAPRKPSPDLRLM 657
Db 591 KOAVGHE---ENAQCKATSDNVIQSSETCSSQSDSTTSA---LPPS-----W 633

QY 658 GAFASLDLKSQRKEELDSRTPAGRGAGTSARGHHSREPLLLGPPAARSLPLP---- 713
Db 634 TSSTSVSQKRRKTKTEPQNE-AVTNVESQNTNRN-----NSTALQLVSRPPDSV 686

QY 714 LEG-KAKRSH 722
Db 687 TEGTPTSKEAH 696

RESULT 5
ID Q9BTNO PRELIMINARY; PRT; 628 AA.
AC Q9BTNO;
DT 01-JUN-2001 (TRENBLrel. 17, Created)
DT 01-JUN-2001 (TRENBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
DE Hypothetical 66.3 kDa protein.
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LUNG;
RA Strausberg R.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC003578; AAH03578.1;
DR InterPro; IPR003961; FN_III.
DR InterPro; IPR003599; IG.
DR InterPro; IPR003598; IG_C2.
DR InterPro; IPR003006; IG_MHC.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR000483; LRR_Cterm.
DR InterPro; IPR000372; LRR_Nterm.
DR Pfam; PF00041; fn3; 1.
DR Pfam; PF00047; ig; 1.
DR Pfam; PF00560; LRR; 6.
DR Pfam; PF01463; LRRCT; 1.
DR PRINTS; PR00019; LEURICHRPT.
DR SMART; SM00060; FN3; 1.
DR SMART; SM00409; IG; 1.
DR SMART; SM00408; IGC2; 1.
DR SMART; SM00082; LRRCT; 1.
DR SMART; SM00013; LRRNT; 1.
DR SMART; SM00369; LRR_TYP; 6.
KW Hypothetical protein; Immunoglobulin domain.
SQ SEQUENCE 628 AA; 66259 MW; FEDC5A5056ABC5FC CRC64;

Query Match 38.4%; Score 1573; DB 4; Length 628;
Best Local Similarity 53.1%; Pred. No. 7.8e-114;
Matches 312; Conservative 88; Mismatches 174; Indels 14; Gaps 5;

```


QY 16 AVDAACPKYCVQNLSESLGTLCPKSKGLLFPVPPDIDRRTVELRLGGNFIIHSRODFANM 75
 Db 23 ATPSPCPRRCRCQOTQSLVSLVPCGAGLLFVPPSLDRRAAELELADNFIAASVRRDLANM 82
 QY 76 TGLVDLTSRNTIHSIQPFSLDLESRLHSDNRLSLGDETLRGLVNLQHLIVNNO 135
 Db 83 TGLLHLSLRNTRIRVAAGAFADLRALRALHLDGNRLTSLGEGQLRGLVNLRLHLLSNNQ 142
 QY 136 LGGIADAEFDELTLLEDDLSYNLHGLPWSVRRMVLHQLSDHNLHLLDHIABGTFAD 195
 Db 143 LAALAGALDCAETLEDDLSYNLLEQLPWEALGRGNVNTLGLDHNLLASVPAGAFSR 202
 QY 196 LOKLARLDLTSNRLOKLPDPIFARSQASALTATPFAPPLS---FSFGNPLHNCCELLW 252
 Db 203 LHKLARLDLTSNRLOKLPDPIFARSQASALTATPFAPPLS---LPLLRPRGSPASALVLAFCGNPLHNCCELLW 259
 QY 253 LRRLEDDLETCGSPGKGRYFHWVREEFVCEPPLITQHTHKLLEGOAATLKCKA 312
 Db 260 LRRLEDDLETCGSPGKGRYFHWVREEFVCEPPLITQHTHKLLEGOAATLKCKA 319
 QY 313 ICDSPHLHWAPDRLVGNSSRTAVYDNGTLDIFITTSQDSGAFTCIAANAAGATAMV 372
 Db 320 VGDPEPRVWSPQGRLLGNSSRARAFNGTLELLVTEPGDGGIFTCIAANAAGATAV 379
 QY 373 EVSI--VOLPHLSNSTRTAPPKSLSDITGSSKTSRGGGSGGGEPPKSPPERAVLVSE 430
 Db 380 ELTVGPPPPQLANSTCDPPDRGDPDALTTPSAASASAKVADTG---PPTDRGVQVTE 435
 QY 431 VTTTSALVWKSVSASAPRVKMYQOYNCSDDDEVLYRMIIPASNKAFVNNVLSVGTGYDLC 490
 Db 436 HGATAALVQWQDREPIGRMYQIQYNSADDILYRMIIPAESRFLTLASGRTYDLC 495
 QY 491 VLAWDDTATTLTATNIVGCAOFFKADYPOCOSMHSOILGTMILVGGIIVATLLVFI 550
 Db 496 VLAVYEDSATGLTATRVGCAARFSTEPALRPGCAHPFLGTMILVGGIIVASVLVFI 555
 QY 551 VILVMRYKVCNHEAP--SKMAAAVSNVYQTNAGQPPPPSSAPAGAPP 596
 Db 556 FVLLMRYKVGQPPGKAKIPAPVSSVCSQINGALGPTTPAPPAPPEP 603

RESULT 6
 ID Q9P244 PRELIMINARY; PRT; 700 AA.
 AC Q9P244
 DT 01-OCT-2000 (TReMBLrel. 15, Created)
 DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
 DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)
 DE KIAA1484 protein (Fragment).
 GN KIAA1484.
 OS Homo sapiens (Human)
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20277482; PubMed=10819331;
 RA Nagase T., Kikuno R., Ishikawa K., Hirosewa M., Ohara O.;
 RT "Prediction of the coding sequences of unidentified human
 genes.XVII.The complete sequences of 100 new cDNA clones from brain
 which code for large proteins in vitro."
 RL DNA Res. 7:143-150(2000).
 DR EMBL: AB040917; BAA96008.1;
 DR InterPro: IPR003961; FN_III.
 DR InterPro: IPR003598; Ig_C2.
 DR InterPro: IPR003006; Ig_MHC.
 DR InterPro: IPR001611; LRR.
 DR InterPro: IPR000483; LRR_Cterm.
 DR InterPro: IPR003592; LRR_out.
 DR InterPro: IPR003591; LRR_typ.
 DR InterPro: IPR001005; Myb_DNA_binding.
 DR Pfam: PF00041; fn3; 1.

DR Pfam: PF00047; ig; 1.
 DR Pfam: PF00560; LRR; 6.
 DR Pfam: PF01463; LRRCT; 1.
 DR SMART: SM00060; FN3; 1.
 DR SMART: SM00408; IGE2; 1.
 DR SMART: SM00370; LRR; 4.
 DR SMART: SM00082; LRRCT; 1.
 DR SMART: SM00369; LRR_TYP; 1.
 KW Immunoglobulin domain.
 FT NON TER
 SQ SEQUENCE 700 AA; 74865 MW; 30D0432242494159 CRC64;
 Query Match 38.4%; Score 1571.5; DB 4; Length 700;
 Best Local Similarity 45.7%; Pred. No. 1.2e-113;
 Matches 352; Conservative 106; Mismatches 203; Indels 109; Gaps 18;
 QY 59 LGGFIHISRODFANMGTGLDLESRLHSDNRLSLGDETLRGLVNLQHLIVNNO 118
 Db 1 LTDNFIARVRRDRFANMTSLVHLTSLRNTIGQVAAAGAFADLRALRALHLDNRLAEVRGD 60
 QY 119 TLRGLVNLQHLIVNNO LGGIADAEFDELTLLEDDLSYNLHGLPWSVRRMVLHQL 178
 Db 61 QLRLGULNLRHLI LGGNQRVRESAFAFLSTVEDLDSYNLLEALPWEAVQOMVNLTL 120
 QY 179 SLDNHLDLHIAEGTFADLQKLARLDLTSNRLOKLPDPIFARSQASALTATPFAP-PLSF 237
 Db 121 TLDHNLHDIHIAEGTFVQLHKLVRDMLTSNRLOKLPDPIFARSQASALTATPFAP-PLSF 176
 QY 238 SFGNPLHNCCELLWRLERLDDLETCGSPGKGRYFHWVREEFVCEPPLIT-QHTH 296
 Db 177 SFGNPLHNCCELLWRLERLDDLETCGSPGKGRYFHWVREEFVCEPPLITQROAGG 236
 QY 297 KLLVLEGOAATLKCKAIGDPSPLHWHVAPDRLVGNSSRTAVYDNGTLDIFITTSODSGA 356
 Db 237 RALVVEGOAVSLRCRAVGDPEPVVHWVAPDORLLGNSSRTRVGGDGLDVTITLRDLSGT 296
 QY 357 FTCTAANAAGSEATAMVEVSIYQLPHLSNSTRTAPPKSLSDITGSSKTSRGGGSGGGE 416
 Db 297 FTCTAANAAGSEATAMVEVSIYQLPHLSNSTRTAPPKSLSDITGSSKTSRGGGSGGGE 349
 QY 417 PPKSPPERAVLVSEVTTTSALVWKSVSASAPRVKMYQOYNCSDDDEVLYRMIIPASNKAF 476
 Db 350 --DSAAERLVAEELTNSVLRWPAQRVPEGRMYQYNSVDDSLVYRMIIPESVQTF 407
 QY 477 VVNLVSGTYDLCVLAWDDTATTLTATNIVGCAOFFKADYPOCOSMHSOILGTMIL 536
 Db 408 LVNDLAAGRAYDLCVLAVIDDGTALPATRVVGVQVFTTAGDPAPCRPRAHFLGTMII 467
 QY 537 VIGGIIVATLLVFIIVLMVRYKVCNHEAPSKMAAAVSNVYQTNAGQPPPPSSAP 591
 Db 468 AIGGVIVASVLVFIIVLMIRYKVVYDGDGSRVVKGRSRLPRVSHVCSQINGA--GTGAQA 524
 QY 592 AGAPP-----QGPVKKVVRNELLDFTLASLARADSSSSSSSSLSGSGEAAGLGRA 638
 Db 525 APALPAODHYREALREVESQAAPAVAVE-----AKAMEAETASA----- 562
 QY 639 PWRIPPSAPRPKPSLDRLMGAFASLDLKSQRKEBELDLSRTPAGRGACTSARGHHSRDREP- 697
 Db 563 -----EPEVVLGRSLGGSATSLCLLPEEETSGBEESR-----AAVGRPRRSRG 604
 QY 698 LLGPPAARARSLPLP-----LEGKAKRSHSFDMDGFAAAAAGVYVPGYSPRKVSNLWT 753
 Db 605 ALEPPTSAPTLALVPGAAARPRPQRYSPD-GDY-----GALFQSHSYPRRARR--T 655
 QY 754 KRSLSVNGMLLPFEESDVLVARG-----TFGSSEWVMESTV 789
 Db 656 KRHRST-----PHLDGAGGGAAGDGLGLGSAACLAFTSTEWMLLESTV 700
 RESULT 7
 ID Q99KT6 PRELIMINARY; PRT; 492 AA.

Q96C50 PRELIMINARY; PRT; 450 AA.
 AC Q96C50;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE Hypothetical 47.4 kDa protein (fragment).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=OVARY;
 RA Strausberg R.;
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC014678; AAH14678.1;
 DR InterPro; IPR003961; FN_III.
 DR InterPro; IPR003006; IG_MHC.
 DR Pfam; PF00041; fn3; 1.
 DR Pfam; PF00047; ig; 1.
 KW Hypothetical protein.
 FT NON_TER 1
 SQ SEQUENCE 450 AA; 47417 MW; 93CEE7B44D2C1FB6 CRC64;
 Query Match 18.9%; Score 775; DB 4; Length 450;
 Best Local Similarity 38.8%; Pred. No. 6.7e-52;
 Matches 188; Conservative 75; Mismatches 138; Indels 84; Gaps 13;

Qy 297 KLLVLEQATLCKKAIKADPSPLIHWVADDDLVGNSSRTAVYDNGTLDIFITTSQDGA 356
 Db 6 RALVVEGQAVSRCAVGDPEVHWVADPDRLLGNSSRVRVGGTLDVITLDRSDG 65
 Qy 357 FTICIAANAAGEATAMVEYSIVOLPHLSNSTRTAPPKRSLSDITGSSKTRGGGGGGE 416
 Db 66 FTICIASNAAGEATAPVEVCVPEPLM-----APPAAPPETEPGSSDIATPRPGAN 118
 Qy 417 PPKSPPERAVLSEVTTTSALVKSAPRVKMYLQYNCSDDEVLIRYRMIIPASNKAF 476
 Db 119 --DAAERLVAELTNSVLRWPAQRPVPGIRMYQVYNSVDDSLVYRMIPTSTQTF 176
 Qy 477 VVNLVSGTGYDLCLVAMWDDTATLITATNIVGCAQFFTKADYPOCOSMHSQILGGTMIL 536
 Db 177 LVNDLAAGRAYDLCVLAVYDDGATALPATRVVGVQVFTTAGDPAFCPLRAHFLGGTMI 236
 Qy 537 VGGIIVATLVFIVILMVRVKV---CNHEAPSKMAAAVSNVYQVYNTGAGQPPPPSSAP 591
 Db 237 ALGGVIVASLVFIVLLMIRYKYGDSRRVKGSRSLPRVSHVCSQTNGA---GTGAHQ 293
 Qy 592 AGAPP-----QGPVKVYVRNELLDFITASLARADSSSSSSSLGSGEAAAGLGRA 638
 Db 294 APALPAQDHYEALREVESQAAPAVAVE-----AKAMEAETASA----- 331
 Qy 639 PWRIPSPAPRPPSILDRMLMGAFASLDLKSQRKEILLDRTPAGRGAGTARSCHHSDRP 697
 Db 332 -----EPEVLGRSLGGSATSLCLLPSSEETSGEER-----AAVGRRRSRG 373
 Qy 698 LIGPPAARARSILPLP----LEGKAKRSHSFDMGDFAAAAAGVWPGGYSPPRKYNIWT 753
 Db 374 ALEPPTSAPPTLALVPGGAARPRFOQRYSD-GDY-----GALFQSHYPRRARR--T 424
 Qy 754 KRSLS 758
 Db 425 KHRHS 429

RESULT 10
 Q96C50 PRELIMINARY; PRT; 450 AA.
 AC Q96C50;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)

DE Hypothetical 33.2 kDa protein.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=EYE;
 RA Strausberg R.;
 RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC014040; AAH14040.1;
 DR EMBL; BC015581; AAH15581.1;
 DR InterPro; IPR003961; FN_III.
 DR Pfam; PF00041; fn3; 1.
 DR SMART; SM00060; FN3; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 324 AA; 33192 MW; C84CCC9A811E877D CRC64;
 Query Match 16.0%; Score 653; DB 4; Length 324;
 Best Local Similarity 34.1%; Pred. No. 1.3e-42;
 Matches 161; Conservative 54; Mismatches 107; Indels 150; Gaps 10;

Qy 320 IHWVAPDRLVGNSSRTAVYDNGTLDIFITTSQDGAFTICIAANAAGEATAMVEYSIVOL 379
 Db 1 MHWVGPDRLVGNSSRVARAFNGLIEIGVTGADGGYTCIATNPAGATARVELRVLAL 60
 Qy 380 PHLNSNSTRTAPPKSRSLSDITGSSKTRGGGGGGEPPKPPPERAVLSEVTTTSALVYK 439
 Db 61 PHGGNSSAEGGRPGP--SDIAASARTAAEGEGTLESEP-----AVQVTEVTATSLVLS 111
 Qy 440 WSVKSAAPRVKMYLQYNCSDDEVLIRYRMIIPASNKAFVNNLVSGTGYDLCLVAMWDDTA 499
 Db 112 WGPCRPADPVMFMFQIYNSSEDETLIRYRMIIPASNHFLKHLVPGADYDLCILLALSPAAG 171
 Qy 500 -TTLTATNIVGCAQFFTKADYPOCOSMHSQILGGTMIIVGIIIVATLLVFTVILMVRVK 558
 Db 172 PSDLTATRLGCAHFFSTLPASPLCHALQAHVLLGGTLTVAVGGVLAALLVFTVALLVGR 231
 Qy 559 -VCNHEAPSKMAAAVSNVYQVYNTGAGQPPPPSSAPAGAPQGPVKVYVRNELLDFITASLAR 617
 Db 232 GAGNGRLPKL---SHVQSQTNGG---PSPTPKAHPRSP----- 266
 Qy 618 ASDSSSSSLGSGEAAAGLGRAPWRIPSPAPRPPKPSLDRMLMGAFASLDLKSQRKEILLDR 677
 Db 267 -----PRP----- 269
 Qy 678 TPAGRGACTSARGHHSDRREPLGLPPAARARSILPLLEGKAKRSHSFDMGDFAAAAAGV 737
 Db 270 -----QRCSLDLGD----- 279
 Qy 738 VPGGYSPPRKYNIWTKRSLVNGMLLPPFEESDLVARGTFCGSSSEWVMESTV 789
 Db 280 -AGCYGARRLLGGAWARSHSVHGGLLG-----AGCRGVGSAERLESVV 324

RESULT 11
 Q96C50 PRELIMINARY; PRT; 640 AA.
 AC Q96C50;
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)

DE KIAA1580 protein (Fragment).
GN KIAA1580.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20450683; PubMed=10997877;
RA Nagase T., Kikuno R., Nakayama M., Hirosewa M., Ohara O.,
RT xviii. The complete sequences of unidentified human genes.
RT code for large proteins in vitro.;
RL DNA Res. 7:273-281(2000).
DR EMBL; AB046800; BAB13406.1; .
DR HSSP; P22888; 1LUT.
DR InterPro; IPR003599; IG.
DR InterPro; IPR003598; IG_c2.
DR InterPro; IPR003600; IG_like.
DR InterPro; IPR003006; IG_MHC.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR000483; LRR_Cterm.
DR InterPro; IPR000372; LRR_Nterm.
DR InterPro; IPR003592; LRR_Out.
DR InterPro; IPR003591; LRR_Cyp.
DR Pfam; PF00047; IG; 1.
DR Pfam; PF00560; LRR; 9.
DR Pfam; PF01463; LRRCT; 1.
DR Pfam; PF01462; LRRNT; 1.
DR PRINTS; PR00019; LEURICHRPT.
DR SMART; SM00409; IG; 1.
DR SMART; SM00408; IGC2; 1.
DR SMART; SM00410; IG_like; 1.
DR SMART; SM00370; LRR; 6.
DR SMART; SM00082; LRRCT; 1.
DR SMART; SM00013; LRRNT; 1.
DR SMART; SM00369; LRR_Typ; 8.
KW Immunoglobulin domain.
FT NON_TER
SQ SEQUENCE 640 AA; 71949 MW; 6058974872636838 CRC64;
Query Match 10.6%; Score 435.5; DB 4; Length 640;
Best Local Similarity 24.5%; Pred. No. 3.1e-25;
Matches 163; Conservative 87; Mismatches 238; Indels 177; Gaps 19;
QY 2 ETLGLGLAFGMFAV----VDACPKYCVCONLSESLGTLCPKGLLFPVPPIDRRIVEL 57
Db DPLLVLLALQLLVVAGLVRAQTCPSCVCSNQFSK--ICVRKNLREVPDGIWTRLL 81
QY 58 RLGNFIHISQDPANMTGLVDTLGRNTISHIQPFSLDLSRLSLHLDNSRPLSGE 117
Db NLHENQIQIKVNSFKHURHLEILQLSNHRTIEIGAFGLANLNTLEFDNRLLTIPN 141
QY 118 DTLRGLVNLQHLIVNHN-----QLGGTADAEAFEDF----- 147
Db 142 GAFVYLSKLELWLRNPNIESIPYAFNRIPSLRRLDLGELKRLSYISEGAFGLSNLRY 201
QY 148 -----LITLEDLDSYNNLHCLPWSVRRMVLNHLSDHNLHLDHIAEG 191
Db 202 LNLAMCNLREIPNLPLIKLDELDDUSGNHLSAIRPGSQGLMHLKLMWIOSQIQVIERN 261
QY 192 TFADLQKARLDLTSNRLOKPPDPIFARSOASALTATFPAPPLSFGGNLHNCCELL 251
Db 262 AFDNLQSLVEINLAHNNLTLPHD-LF-----TPLLHLERILHLHHPMNCNDIL 310
QY 252 WLRRLERD-----DDLETCSGPGGLKGRYFHWVREBEFVCPPLITQHTKLLVLEGOA 305
Db 311 WLSWIKDMAPNSNTACCARCNTPPNLKGRIYIGELDQNYFTCVAPVIVEPPADLNVTEGMA 370
QY 306 ATLKKAIGDPSPLIHVAPDRDLVGNSS---RTAVYDNGLIDFITTSQDSGAPFTAA 362
Db 371 AELKCR-A-STSLTSVSWITPENGVTMTHGAYKVRIVASDGLTFLNFTVQDTGMVTCMS 429

363 NAAGEATAMVEYSIVQLPHLSNSTSRITAPPKSRSLSDITGSSKTRGGGGGGGEBPKPSP 422
430 NSVGNNTASATL-----NVTAATTPFSYFSTVT----- 458
423 ERAVLVSEVHTTSALVKSWSKAPRVKMYQLQYNCSDDEVLIYRMIIPASKAFVNNLV 482
459 -----VETMFSQDEARTDNNV 476
483 SGTGYDLCLVLANWDDT--ATLLTATNIVGCAQFFT-----KADYPOCSMHSQILGGTM 534
477 GPT-----PVDWETINVTSLTPQSTRSTKTFPIPTVDINSQIPGI-----DEVKTIK 527
535 ILVIGIIVATLLVFIIVLMVRYK-----CNHEAPSKMAAAVSNVYVQTMGAOP----- 584
528 III--GCFVAITLMAAVMLVIFYKMKRKHQHRONHHPTR-TVEIINVDDEITGDTPMESH 584
585 -PPPS 588
585 LPMPA 589
RESULT 12
Q9BZ20 PRELIMINARY; PRT; 606 AA.
AC Q9BZ20
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE BA438B23.1 (Neuronal leucine-rich repeat protein) (CDNA FLJ131810 fis,
DE clone NT2R12009289, weakly similar to carboxypeptidase N 83 kDa
DE chain).
GN BA438B23.1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Babbage A.;
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Ishibaashi T., Kanehori K., Yosida M., Watanabe S., Ishida S., Ono Y.,
RA Hotuta T., Hiraoka S., Murakawa K., Takiguchi S., Kusano J.,
RA Watanabe M., Fujimori K., Tanai H., Ishida M., Yamashita H., Chiba Y.,
RA Sugiyama T., Irie R., Otsuki T., Sato H., Wakamatsu A., Ishii S.,
RA Yamamoto J., Isono Y., Kawai-Hio Y., Saito K., Nishikawa T.,
RA Kimura K., Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K.,
RA Wagatsuma M., Takahashi-Fujii A., Oshima A., Sugiyama A., Kawakami B.,
RA Suzuki Y., Sugano S., Nagahari K., Masuho Y., Nagai K., Isogai T.;
RA "NEDO human cDNA sequencing Project";
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
RL EMBL; AL353746; CAC22713.1; .
DR InterPro; IPR003599; IG.
DR InterPro; IPR003598; IG_c2.
DR InterPro; IPR003600; IG_like.
DR InterPro; IPR003006; IG_MHC.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR000483; LRR_Cterm.
DR InterPro; IPR000372; LRR_Nterm.
DR InterPro; IPR003591; LRR_Typ.
DR Pfam; PF00047; IG; 1.
DR Pfam; PF00560; LRR; 10.
DR Pfam; PF01463; LRRCT; 1.
DR PRINTS; PR00019; LEURICHRPT.
DR SMART; SM00409; IG; 1.
DR SMART; SM00408; IGC2; 1.
DR SMART; SM00410; IG_like; 1.
DR SMART; SM00082; LRRCT; 1.
DR SMART; SM00013; LRRNT; 1.
DR SMART; SM00369; LRR_Typ; 10.

```

KW Immunoglobulin domain.
SQ SEQUENCE 606 AA; 68065 MW; CB608E281B066B9D CRC64;
  Query Match          10.2%; Score 418.5; DB 4; Length 606;
  Best Local Similarity 26.0%; Pred. No. 6.1e-24;
  Matches 125; Conservative 67; Mismatches 157; Indels 131; Gaps 12;
  QY 21 CPKVCVQNLESGLTLCPSKGLLVFPDPIDRRRTVELRLGNGNFIHHSRQDFANMTGLVD 80
  DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
  Db CPVCSNPFQSKV-VCTRRGLSEVPOGIPSNTRYLNLMENNIIQMIQADTFRHLHLLEV 103
  QY 81 LTLSRNTTISHQPPFDLESRLSRLSDNSRPLSGEDTDLRGLVNLQHLIVNNOLGGIA 140
  DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
  Db LQLGRNSIRQIEVGFAGNFASLNTLELFDNMLTIPSGAFEYLSKLRELWLRNPPIESIP 163
  QY 141 DEAFEDF-----DFFLLLEDLDLSYNNLHGHPWDSVRRMVNLH 176
  DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
  Db SYAFNRVPMLRDLGELKLEYLSEGAFFGFLNLKYLNLGMCNLDKMP--NLPLVGLGE 221
  QY 177 QLSLDHLLDHIAGTFFADLQKLAR-----LDLTSNRLOKLPDPP 216
  DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
  Db ELEMGNHFPFIRPGSFHGLSSLKLWYMNHGHERNAFDGLASLVELNLAHNNLSLPHD- 280
  QY 217 IFARQASALATATFPAPPLSFSGNPLHCELLWLRLERD---DLETGCG--SFGG 270
  DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
  Db TFLRVLVELLHHLHPNMCDCDILMLAWMLREYIPTNSTCCGRCHARPH 330
  QY 271 LKGRVFWHVEEFCVPEPLITQHTKLLVLEGOAATLCKKCAIGDPPSLIHWVAPDDRLV 330
  DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
  Db MRGRYLVEVDOAFQCSAPFIMDPRDLNISSEGRMAELKCRF--PPMSVKWLLPENGTVL 388
  QY 331 GNS---RTAYVNGTLDFITTSODSGAFTCIAAAG--EATAMVEVSIQVLPHLGNS 385
  DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
  Db SHASRHPRI SVLNDGTLNFSHVLLSDTGVTYTCMTVNVAGNSAYLNVAEL-NTSNY 447
  QY 386 TSRTAPPKRSLSDITGSKTSRGGGGGGGGPPKPPERAVLVSEVTTTSALVKKVSVSKS 445
  DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
  Db SFFT-----TVTVETTEIS----- 461
  QY 446 APRVWVQLYQNCSDDEVLIIYRMI PASNKAFVWNNLVSGTVGDLVCLVLAEMDDTTATL 505
  DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
  Db -----PEDITRKYKPVFTS-----TGY-----OPATVTTSTT 488
  QY 506 NIVGCAQPFYTKADYK-----QCQSMHSQILGTMILVIGGIIIVATLLVFIIVLVRYKV 559
  DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
  Db VLIQTRTPVKQVAVPATDITDKMQLTSLDEVMTTKIII--GCFVAVTLLAAAMLIVFYKL 546
  QY 560 -CNHEAPSKMAAAVNSVYQNGAQPPPPPSAPAGAPPOG 598
  DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
  Db KRKHQRSTVTAARTVEIIQVD-EDIPATAAATAAPSG 585

RESULT 14
Q8VI35
ID Q8VI35; PRELIMINARY; PRT; 648 AA.
AC Q8VI35;
DT 01-MAR-2002 (TRENBLrel. 20, Created)
DT 01-MAR-2002 (TRENBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)
DE Brain tumor-associated protein MBAG1.
OS Mus musculus (Mouse)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/J;
RA Wang J., Bin L., Li G.;
RL "Brain-specific gene, downregulated in brain tumor."
RL Submitted (AUG-2000) to the EMBL/GenBank/DBDJ databases.
DR EMBL; AF300458; AAL67671.1;
DR InterPro; IPR003598; IG_c2.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR000483; LRR_Cterm.
DR InterPro; IPR000372; LRR_Nterm.
DR InterPro; IPR003592; LRR_out.
DR InterPro; IPR003591; LRR_typ.

```

```

KW Immunoglobulin domain.
SQ SEQUENCE 606 AA; 68065 MW; CB608E281B066B9D CRC64;
  Query Match          10.2%; Score 418.5; DB 4; Length 606;
  Best Local Similarity 26.0%; Pred. No. 6.1e-24;
  Matches 125; Conservative 67; Mismatches 157; Indels 131; Gaps 12;
  QY 21 CPKVCVQNLESGLTLCPSKGLLVFPDPIDRRRTVELRLGNGNFIHHSRQDFANMTGLVD 80
  DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
  Db CPVCSNPFQSKV-VCTRRGLSEVPOGIPSNTRYLNLMENNIIQMIQADTFRHLHLLEV 103
  QY 81 LTLSRNTTISHQPPFDLESRLSRLSDNSRPLSGEDTDLRGLVNLQHLIVNNOLGGIA 140
  DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
  Db LQLGRNSIRQIEVGFAGNFASLNTLELFDNMLTIPSGAFEYLSKLRELWLRNPPIESIP 163
  QY 141 DEAFEDF-----DFFLLLEDLDLSYNNLHGHPWDSVRRMVNLH 176
  DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
  Db SYAFNRVPMLRDLGELKLEYLSEGAFFGFLNLKYLNLGMCNLDKMP--NLPLVGLGE 221
  QY 177 QLSLDHLLDHIAGTFFADLQKLAR-----LDLTSNRLOKLPDPP 216
  DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
  Db ELEMGNHFPFIRPGSFHGLSSLKLWYMNHGHERNAFDGLASLVELNLAHNNLSLPHD- 280
  QY 217 IFARQASALATATFPAPPLSFSGNPLHCELLWLRLERD---DLETGCG--SFGG 270
  DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : ~~~
  Db TFLRVLVELLHHLHPNMCDCDILMLAWMLREYIPTNSTCCGRCHARPH 330
  QY 271 LKGRVFWHVEEFCVPEPLITQHTKLLVLEGOAATLCKKCAIGDPPSLIHWVAPDDRLV 330
  DB : : : : ~~~ : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
  Db MRGRYLVEVDOAFQCSAPFIMDPRDLNISSEGRMAELKCRF--PPMSVKWLLPENGTVL 388
  QY 331 GNS---RTAYVNGTLDFITTSODSGAFTCIAAAG--EATAMVEVSIQVLPHLGNS 385
  DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : ~~~
  Db SHASRHPRI SVLNDGTLNFSHVLLSDTGVTYTCMTVNVAGNSAYLNVAEL-NTSNY 447
  QY 386 TSRTAPPKRSLSDITGSKTSRGGGGGGGGPPKPPERAVLVSEVTTTSALVKKVSVSKS 445
  DB : : : : ~~~ : : : : ~~~ : : : : ~~~ : : : : ~~~ : : : : ~~~ : : : : ~~~
  Db SFFT-----TVTVETTEIS----- 461
  QY 446 APRVWVQLYQNCSDDEVLIIYRMI PASNKAFVWNNLVSGTVGDLVCLVLAEMDDTTATL 505
  DB : : : : ~~~ : : : : ~~~ : : : : ~~~ : : : : ~~~ : : : : ~~~ : : : : ~~~
  Db -----PEDITRKYKPVFTS-----TGY-----OPATVTTSTT 488
  QY 506 NIVGCAQPFYTKADYK-----QCQSMHSQILGTMILVIGGIIIVATLLVFIIVLVRYKV 559
  DB : : : : ~~~ : : : : ~~~ : : : : ~~~ : : : : ~~~ : : : : ~~~ : : : : ~~~
  Db VLIQTRTPVKQVAVPATDITDKMQLTSLDEVMTTKIII--GCFVAVTLLAAAMLIVFYKL 546
  QY 560 -CNHEAPSKMAAAVNSVYQNGAQPPPPPSAPAGAPPOG 598
  DB : : : : ~~~ : : : : ~~~ : : : : ~~~ : : : : ~~~ : : : : ~~~ : : : : ~~~
  Db KRKHQRSTVTAARTVEIIQVD-EDIPATAAATAAPSG 585

RESULT 13
Q96A85
ID Q96A85; PRELIMINARY; PRT; 649 AA.
AC Q96A85;
DT 01-DEC-2001 (TRENBLrel. 19, Created)
DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TRENBLrel. 20, Last annotation update)
DE Hypothetical 72.3 kDa protein.
GN BAG.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6;
RA Wang J.;
RL Thesis (2000); Zhongshan Medical University, Guangzhou, China.
RL EMBL; AJ297858; CA82651.1;
DR InterPro; IPR003006; IG_MHC.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR000483; LRR_Cterm.
DR InterPro; IPR000372; LRR_Nterm.
DR Pfam; PF00047; Ig; 1.
DR Pfam; PF00560; LRRCT; 8.
DR Pfam; PF01463; LRRCT; 1.
DR Pfam; PF01462; LRRNT; 1.
DR SMART; SM00013; LRRNT; 1.
DR Hypothetical protein; Signal.
FT SIGNAL 1 38 POTENTIAL.
SQ SEQUENCE 649 AA; 72313 MW; 9C11C5ABC7E536CF CRC64;
  Query Match          9.4%; Score 385; DB 4; Length 649;
  Best Local Similarity 23.8%; Pred. No. 2.7e-21;

```

THIS PAGE BLANK (USPTO)

